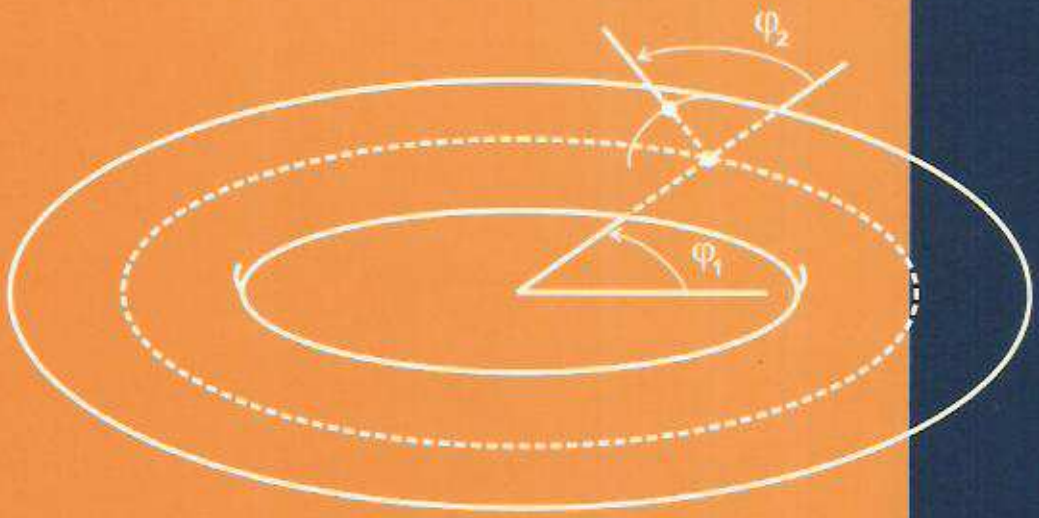
MECÂNICA ANALÍTICA

Nivaldo A. Lemos

2ª edição



Livraria da Física
Editora

A compreensão das teorias físicas contemporâneas exige um domínio das idéias e dos métodos da mecânica analítica, cuja importância intrínseca foi realçada nas últimas décadas do século XX pelo rápido desenvolvimento das teorias do caos e dos sistemas dinâmicos. Neste livro, destinado primordialmente a cursos de graduação, os formalismos lagrangiano e hamiltoniano são introduzidos de forma clara, com ênfase nas propriedades de simetria e invariância e nos aspectos estruturais da mecânica, de importância fundamental para a Física moderna.

Numerosos exemplos tornam o texto acessível aos estudantes típicos. O formalismo é aplicado à dinâmica do corpo rígido, à teoria das pequenas oscilações e à dinâmica relativística. As transformações canônicas e o método de Hamilton-Jacobi são tratados com simplicidade e precisão. Discute-se, ainda, a extensão dos formalismos de Lagrange e de Hamilton a sistemas contínuos, descritos por campos. Além da exposição ampla e detalha-

MECÂNICA ANALÍTICA

Editora Livraria da Física

Nivaldo A. Lemos

MECÂNICA ANALÍTICA

Editora Livraria da Física
São Paulo – 2007 – 2ª edição

Copyright 2007: Editora Livraria da Física

Editor : José Roberto Marinho
Capa : Arte Ativa
Revisão : Osvaldo Pessoa Jr.
Impressão : Gráfica Paym

Dados de catalogação na Publicação (CIP)Internacional
(Câmara Brasileira do Livro, SP, Brasil)

Lemos, Nivaldo A.
Mecânica analítica / Nivaldo A. Lemos. --
2. ed. -- São Paulo : Editora Livraria da Física, 2007.

Bibliografia.

1 . Mecânica I . Título.

04 – 0173 CDD - 531

Índices para catálogo sistemático:

1. Física quântica 531

ISBN : 85-88325-24-1

Editora Livraria da Física

Telefone : 0xx11 – 3816 7599
Fax : 0xx11 – 3815 8688
e-mail : livraria@if.usp.br
Página na internet : www.livrariadafisica.com.br

MECÂNICA ANALÍTICA

Nivaldo A. Lemos

Departamento de Física
Universidade Federal Fluminense

Dedicado às minhas filhas

Cintia, Luiza e Beatriz

Physical laws should have mathematical beauty.

Paul Adrien Maurice Dirac

Poets say science takes away from the beauty of the stars – mere globs of gas atoms. I too can see the stars on a desert night, and feel them. But do I see less or more? The vastness of the heavens stretches my imagination. A vast pattern – of which I am a part. It does not do harm to the mystery to know a little about it. Far more marvellous is the truth than any artists of the past imagined it. What men are poets who can speak of Jupiter if he were a man, but if he is an immense spinning sphere of methane and ammonia must be silent?

Richard Phillips Feynman

PREFÁCIO

Lend me your ears and I'll sing you a song,
And I'll try not to sing out of key.
With a Little Help From My Friends
John Lennon & Paul McCartney

Pode-se afirmar, sem exagero, que a mecânica analítica é o alicerce da física teórica. O grandioso edifício da teoria quântica foi erigido sobre a base da mecânica analítica, particularmente na forma devida a Hamilton. A mecânica estatística e as teorias de campos das partículas elementares são fortemente marcadas por elementos estruturais extraídos da mecânica clássica. Ademais, o vertiginoso desenvolvimento das teorias do caos e dos sistemas dinâmicos em geral promoveu um renascimento da mecânica clássica nas últimas décadas do século XX. Assim, o estudo de praticamente qualquer ramo da física atual requer uma formação sólida em mecânica analítica, a qual, por si só, continua sendo de enorme importância por suas aplicações em engenharia e na mecânica celeste.

Ao longo de mais de duas décadas, de forma intermitente, o autor vem ministrando a disciplina de mecânica analítica no curso de graduação em Física da Universidade Federal Fluminense. O presente texto, fruto dessa experiência e de uma atração ininterrupta pelo assunto, destina-se a estudantes de graduação que tenham passado por um curso intermediário de mecânica num nível comparável ao de Symon (1971) ou Marion & Thornton (1995). Quanto aos pré-requisitos matemáticos, os cursos usuais de cálculo de uma e várias variáveis, equações diferenciais ordinárias e álgebra linear são suficientes.

A mecânica analítica é uma disciplina de caráter eminentemente matemático. Assim, sem descurar dos aspectos físicos, procuramos manter um padrão razoável de rigor matemático na exposição. A maior concessão nesse campo é o emprego de quantidades infinitesimais, que consideramos pedagogicamente aconselhável numa primeira apresentação ao estudante das noções de deslocamento virtual e de famílias contínuas de transformações. Evitamos sobrecarregar o texto com resultados matemáticos auxiliares, exceto quando pareceu possível integrá-los com certa naturalidade ao desenvolvimento do formalismo. Alguns teoremas, de uso mais geral e freqüente, estão provados nos apêndices; nos demais casos, referimo-nos a textos matemáticos onde as demonstrações precisas podem ser encontradas. Entendemos que a progressiva substituição de argumentos heurísticos por procedimentos matematicamente rigorosos está em sintonia com as tendências da física teórica atual, cuja linguagem matemática vem se tornando crescentemente sofisticada e rigorosa.

O aparato matemático da mecânica analítica é muito rico e permite o primeiro contato do estudante com técnicas e conceitos largamente empregados nos mais variados ramos da Física, porém no contexto de uma teoria clássica, no qual a intuição é um guia relativamente seguro. Noções como as de operador linear, autovalor, autovetor, grupo e álgebra de Lie surgem naturalmente, aplicadas a situações mais fáceis de visualizar. O alto grau de generalidade do formalismo da mecânica analítica serve para desenvolver no estudante a capacidade de abstração, tão necessária para tornar possível compreender as teorias físicas contemporâneas. Como observa V. I. Arnold, numerosas teorias matemáticas modernas, com as quais estão associados alguns dos maiores nomes da história da Matemática, devem sua existência a problemas de mecânica. Métodos sofisticadas de geometria diferencial e topologia estão no cerne dos importantes desenvolvimentos recentes, o que inviabiliza uma exposição que lhes faça justiça num curso de graduação. A abordagem tradicional ainda permanece a mais indicada para um primeiro curso de mecânica analítica. Uma vez adquirida a formação básica, o leitor estará apto a enveredar pelos caminhos mais

difíceis da dinâmica qualitativa (Arnold 1976; Thirring 1997). Os livros recentes de Scheck (1994) e de José & Saletan (1998) possibilitam uma transição mais suave do tratamento convencional à formulação altamente matematizada da teoria dos sistemas dinâmicos em linguagem geométrica, essencial para a análise de problemas de estabilidade e sistemas caóticos.

A mecânica analítica é muito mais do que uma mera reformulação da mecânica newtoniana, e os seus métodos não foram concebidos primordialmente para facilitar a resolução de problemas mecânicos específicos. Tentamos deixar isso claro pondo em relevo as propriedades de simetria e invariância e os aspectos estruturais da mecânica. Os livros de Goldstein (1980) e Landau & Lifchitz (1966), nos quais o autor aprendeu mecânica analítica, exerceram grande influência na elaboração do texto.

Nossa exposição caracteriza-se por um grande número de exemplos resolvidos, objetivando torná-la acessível aos estudantes típicos. Ao mesmo tempo, pretendemos que o texto seja estimulante e desafiador aos melhores estudantes. A bibliografia relativamente extensa, com menção a livros avançados e a artigos publicados em periódicos, vários deles bastante recentes, tem por fito excitar a curiosidade do leitor e deixar claro que a mecânica clássica não é peça de museu nem constitui um capítulo encerrado da Física. Pelo contrário, o seu poder de encantamento não parece diminuir com a passagem do tempo e ainda há muitos problemas dignos de investigação, alguns deles monumentais, como o problema da estabilidade do sistema solar, o qual permanece sem resposta conclusiva. Com o intuito de fazer contato com aplicações modernas, resolvemos incluir uma introdução aos sistemas hamiltonianos com vínculos em virtude de sua grande importância na física teórica atual, apesar de ser um tópico com um grau de dificuldade acima da média. Nesse mesmo espírito, fazemos uma breve introdução aos sistemas integráveis e ao teorema de Kolmogorov-Arnold-Moser (KAM), bem como discutimos com alguma profundidade o tratamento do tempo como variável canônica e as teorias com tempo parametrizado, o que se justifica pela amplas aplicações na teoria quântica da gravitação em geral e na cosmologia quântica em particular.

Os exercícios salpicados ao longo dos capítulos são parte inseparável da exposição, e todo estudante sério deve procurar resolvê-los ao primeiro encontro. Os problemas no fim de cada capítulo servem, em sua maioria, como ilustrações e extensões da teoria exposta no texto principal. Há certos problemas, no entanto, que introduzem idéias ou técnicas novas, algumas delas frutos de trabalhos de pesquisa relativamente recentes. Os resultados de alguns problemas são empregados em exemplos, problemas ou no corpo do texto de capítulos subseqüentes.

Pelo estímulo ou de outras maneiras, diversos colegas muito contribuíram para a concretização deste livro, destacadamente Rita Zorzenon dos Santos, Mauro Copelli, Jürgen Stilck e Mucio A. Continentino. Agradecimentos especiais vão para o estudante Bernardo Coutinho Camilo dos Santos e o professor Rubens L. P. G. do Amaral, que contribuíram gentilmente para reduzir o número de erros e obscuridades do texto. Naturalmente, o autor é o único responsável pelos erros remanescentes. Críticas e sugestões são bem-vindas, e podem ser enviadas para o seguinte endereço eletrônico: nivaldo@if.uff.br.

Se, além dos conhecimentos básicos, conseguirmos transmitir ao leitor um pouco do que há de fascinante neste belo ramo da Física e, assim, estimulá-lo a prosseguir trilhando caminhos mais elevados, nosso objetivo terá sido plenamente atingido.

Niterói, novembro de 2003 Nivaldo A. Lemos

Sumário

1 DINÂMICA LAGRANGIANA 5
- 1.1 Princípios da Mecânica Newtoniana . . . 5
- 1.2 Vínculos . . . 11
- 1.3 Princípio de d'Alembert . . . 14
- 1.4 Coordenadas Generalizadas e Equações de Lagrange . . . 19
- 1.5 Aplicações das Equações de Lagrange . . . 26
- 1.6 Potenciais Generalizados e Função de Dissipação . . . 29
- 1.7 Forças Centrais e Teorema de Bertrand . . . 33
- Problemas . . . 39

2 PRINCÍPIO VARIACIONAL DE HAMILTON 43
- 2.1 Rudimentos do Cálculo das Variações . . . 43
- 2.2 Notação Variacional . . . 50
- 2.3 Princípio de Hamilton e Equações de Lagrange . . . 50
- 2.4 Princípio de Hamilton no Caso Não-Holônomo . . . 54
- 2.5 Propriedades de Simetria e Leis de Conservação . . . 63
- 2.6 Conservação da Energia . . . 68
- 2.7 Teorema de Noether . . . 71
- Problemas . . . 76

3 CINEMÁTICA DA ROTAÇÃO 83
- 3.1 Transformações Ortogonais . . . 83
- 3.2 Deslocamentos Possíveis de um Corpo Rígido . . . 89
- 3.3 Ângulos de Euler . . . 91
- 3.4 Rotações Infinitesimais e Velocidade Angular . . . 93
- 3.5 Grupo de Rotações e Geradores Infinitesimais . . . 99
- 3.6 Dinâmica em Referenciais Não-Inerciais . . . 101
- Problemas . . . 108

4 DINÂMICA DO CORPO RÍGIDO 111
- 4.1 Momento Angular e Tensor de Inércia . . . 111
- 4.2 Interlúdio Matemático: Tensores e Diádicos . . . 113
- 4.3 Momentos e Produtos de Inércia . . . 118
- 4.4 Energia Cinética e Teorema dos Eixos Paralelos . . . 119
- 4.5 Diagonalização do Tensor de Inércia . . . 122
- 4.6 Simetrias e Eixos Principais de Inércia . . . 126
- 4.7 Moeda Rolante . . . 128

4.8 Equações de Euler e Rotação Livre 130
4.9 Pião Simétrico com um Ponto Fixo 138
Problemas 145

5 PEQUENAS OSCILAÇÕES **149**
5.1 Caso Unidimensional 149
5.2 Movimento Estacionário e Pequenas Oscilações 154
5.3 Pequenas Oscilações: Caso Geral 157
5.4 Modos Normais de Vibração 158
5.5 Coordenadas Normais 163
5.6 Suplemento Matemático 170
Problemas 173

6 MECÂNICA RELATIVÍSTICA **177**
6.1 Transformações de Lorentz 177
6.2 Cone de Luz e Causalidade 183
6.3 Vetores e Tensores 186
6.4 Campos Tensoriais 189
6.5 Leis Físicas em Forma Covariante 191
6.6 Dinâmica Relativística 194
6.7 Colisões Relativísticas 200
6.8 Dinâmica Relativística na Forma Lagrangiana 203
6.9 Ação à Distância na Relatividade Especial 205
Problemas 208

7 DINÂMICA HAMILTONIANA **213**
7.1 As Equações Canônicas de Hamilton 213
7.2 Simetrias e Leis de Conservação 217
7.3 Teorema do Virial 219
7.4 Formulação Hamiltoniana Relativística 222
7.5 Forma Variacional das Equações de Hamilton 224
7.6 O Tempo Como Variável Canônica 226
7.7 Princípio de Maupertuis 231
Problemas 235

8 TRANSFORMAÇÕES CANÔNICAS **239**
8.1 Transformações Canônicas e Funções Geradoras 239
8.2 Canonicidade e Parênteses de Lagrange 245
8.3 Notação Simplética 247
8.4 Parênteses de Poisson 250
8.5 Transformações Canônicas Infinitesimais 253
8.6 Parênteses de Poisson do Momento Angular 259
8.7 Teoremas de Liouville e de Poincaré 260
8.8 Sistemas Hamiltonianos Vinculados 264
Problemas 275

9 TEORIA DE HAMILTON-JACOBI **281**
9.1 A Equação de Hamilton-Jacobi 281
9.2 Exemplos Unidimensionais 284
9.3 Separação de Variáveis 287
9.4 A Ação Como Função das Coordenadas 292
9.5 Variáveis de Ação e Ângulo 295
9.6 Sistemas Integráveis e Teorema KAM 300
9.7 Variáveis de Ação no Problema de Kepler 306
9.8 Invariantes Adiabáticos 308
9.9 Teoria de Hamilton-Jacobi e Mecânica Quântica 310
Problemas 315

10 TEORIA CLÁSSICA DE CAMPOS **319**
10.1 Teoria de Campos na Forma Lagrangiana 319
10.2 Teorias de Campos Relativísticas 325
10.3 Derivadas Funcionais 327
10.4 Teoria de Campos na Forma Hamiltoniana 330
10.5 Simetrias da Ação e Teorema de Noether 333
10.6 Ondas Solitárias e Sólitons 337
10.7 Campos Vinculados 341
Problemas 346

APÊNDICES **351**

A Notação Indicial **351**

B Funções Homogêneas e Teorema de Euler **357**

C Espaços Vetoriais e Operadores Lineares **359**
C.1 Operadores Lineares 359
C.2 Representação Matricial de Operadores 360
C.3 Autovalores e Autovetores 362
C.4 Produto Interno e Bases Ortonormais 363
C.5 Complemento Ortogonal e Soma Direta 365
C.6 Adjunto de um Operador Linear 366
C.7 Operadores Unitários e Auto-Adjuntos 367

D Diferenciais Exatas **373**
Bibliografia 375
Índice Remissivo 381

Capítulo 1

DINÂMICA LAGRANGIANA

> Lagrange has perhaps done more than any other analyst ... by showing that the most varied consequences respecting the motion of systems of bodies may be derived from one radical formula; the beauty of the method so suiting the dignity of the results, as to make of his great work a kind of scientific poem.
>
> William Rowan Hamilton

Sistemas mecânicos sujeitos a restrições (vínculos) de natureza geométrica ou cinemática ocorrem com freqüência. Em tais situações a formulação newtoniana da dinâmica revela-se inconveniente e antieconômica, pois exige o uso de variáveis redundantes e as forças de vínculo aparecem de forma explícita. O poderoso e elegante formalismo desenvolvido por Lagrange permite escrever as equações de movimento da maioria dos sistemas físicos relevantes a partir de uma única função escalar expressa em termos de coordenadas independentes arbitrárias, com a vantagem adicional de não envolver as forças de vínculo.

1.1 Princípios da Mecânica Newtoniana

A justa apreciação do significado e da abrangência das formulações gerais da mecânica clássica requer uma breve revisão dos princípios fundamentais da mecânica newtoniana, com a qual supõe-se que o leitor já esteja familiarizado. Praticamente desde a sua formulação nos *Principia*, as três leis do movimento de Newton vêm sendo alvo de controvérsias quanto ao seu conteúdo físico e consistência lógica, o que tem gerado propostas de reformulação da versão tradicional com o intuito de escapar às críticas (Eisenbud 1958; Weinstock 1961). Embora a primeira e segunda leis sejam às vezes interpretadas como uma definição de força (Marion & Thornton 1995), adotaremos o ponto de vista que julgamos mais correto, segundo o qual elas são leis genuínas e não meras definições (Anderson 1990). Uma análise detalhada dos aspectos físicos e da estrutura lógica das leis de Newton transcende o escopo desta seção, cujo propósito primordial é servir de referência para o restante da exposição.

Leis do Movimento

Os postulados enunciados a seguir equivalem às três leis do movimento de Newton, mas procuram evitar certas dificuldades lógicas da proposição original.

PRIMEIRA LEI. Existem sistemas de referência, ditos *inerciais*, em relação aos quais toda partícula isolada descreve um movimento retilíneo uniforme.

A existência de um referencial inercial implica a existência de uma infinidade de outros, todos movendo-se entre si em linha reta com velocidade constante. Neste postulado está implícita a noção newtoniana de tempo absoluto, que "flui uniformemente sem relação com qualquer coisa externa" e é o mesmo em todos os referenciais enerciais. Considera-se "isolada" uma partícula suficientemente afastada de todos os objetos materiais.

SEGUNDA LEI. Em qualquer referencial inercial o movimento de uma partícula é regido pela equação

$$m\,\mathbf{a} = \mathbf{F}\,, \tag{1.1.1}$$

onde $\mathbf{a}$ é a aceleração da partícula, m sua massa e $\mathbf{F}$ a força total a que ela está sujeita.

Este postulado pressupõe, implicitamente, que a cada partícula está associada uma constante positiva m, denominada massa, que é a mesma em todos os referenciais inerciais.

TERCEIRA LEI. A cada ação corresponde uma reação igual e oposta, isto é, se $\mathbf{F}_{ij}$ é a força sobre a partícula i exercida pela partícula j, então

$$\mathbf{F}_{ij} = -\mathbf{F}_{ji}\,. \tag{1.1.2}$$

Esta é a lei da ação e reação em sua forma fraca. Em sua versão forte, esta lei declara que, além de iguais e opostas, as forças são dirigidas ao longo da linha que une as partículas; em outras palavras, duas partículas só podem se atrair ou repelir. A terceira lei não tem validade geral, pois as forças entre cargas elétricas em movimento geralmente a violam. Isto deve-se à velocidade finita de propagação das interações eletromagnéticas, o que exige introduzir a noção de campo eletromagnético como mediador de tais interações.

No caso de um sistema contendo várias partículas, supõe-se que a força sobre cada uma delas pode ser decomposta em *forças externas*, produzidas por fontes exteriores ao sistema, e *forças internas*, que devem-se às demais partículas do sistema.[1] Assim, a equação de movimento da i-ésima partícula de um sistema de N partículas é, conforme a segunda lei,

$$\frac{d\mathbf{p}_i}{dt} = \sum_{\substack{j=1\\ j\neq i}}^{N} \mathbf{F}_{ij} \;+\; \mathbf{F}_i^{(e)}\,, \tag{1.1.3}$$

onde

$$\mathbf{p}_i = m_i\,\mathbf{v}_i = m_i\,\frac{d\mathbf{r}_i}{dt} \tag{1.1.4}$$

é o *momento linear* da i-ésima partícula, m_i sua massa, $\mathbf{r}_i$ seu vetor posição, $\mathbf{v}_i$ sua velocidade e $\mathbf{F}_i^{(e)}$ denota a força externa sobre ela.

[1] Descarta-se a possibilidade de uma partícula agir sobre si mesma.

Teoremas de Conservação

Somando as Eqs.(1.1.3) sobre todas as partículas deduz-se[2]

$$\sum_i m_i \frac{d^2\mathbf{r}_i}{dt^2} = \sum_{\substack{i,j \\ i\neq j}} \mathbf{F}_{ij} + \sum_i \mathbf{F}_i^{(e)} = \sum_i \mathbf{F}_i^{(e)} \tag{1.1.5}$$

porque (vide Eq.(A.9) do Apêndice A)

$$\sum_{\substack{i,j \\ i\neq j}} \mathbf{F}_{ij} = \frac{1}{2}\sum_{\substack{i,j \\ i\neq j}} (\mathbf{F}_{ij} + \mathbf{F}_{ji}) = 0 \tag{1.1.6}$$

em virtude da Eq.(1.1.2). Definindo o vetor posição do centro de massa por

$$\mathbf{R} = \frac{\sum_i m_i \mathbf{r}_i}{\sum_i m_i} \equiv \frac{\sum_i m_i \mathbf{r}_i}{M} , \tag{1.1.7}$$

a Eq.(1.1.5) assume a forma

$$M\frac{d^2\mathbf{R}}{dt^2} = \sum_i \mathbf{F}_i^{(e)} \equiv \mathbf{F}^{(e)} , \tag{1.1.8a}$$

ou

$$\frac{d\mathbf{P}}{dt} = \mathbf{F}^{(e)} \tag{1.1.8b}$$

em termos do momento linear total do sistema definido por

$$\mathbf{P} = \sum_i m_i \mathbf{v}_i = \sum_i m_i \frac{d\mathbf{r}_i}{dt} = M\frac{d\mathbf{R}}{dt} . \tag{1.1.9}$$

Infere-se, assim, uma importante lei de conservação.

Teorema da Conservação do Momento Linear. Se a força externa total é zero, o momento linear total de um sistema de partículas é conservado.

O momento angular total do sistema em relação a um ponto Q com vetor posição $\mathbf{r}_Q$ é dado por[3]

$$\mathbf{L}_Q = \sum_i m_i(\mathbf{r}_i - \mathbf{r}_Q) \times (\dot{\mathbf{r}}_i - \dot{\mathbf{r}}_Q) , \tag{1.1.10}$$

onde $\mathbf{r}_i^{(Q)} = \mathbf{r}_i - \mathbf{r}_Q$ e $\mathbf{v}_i^{(Q)} = \dot{\mathbf{r}}_i - \dot{\mathbf{r}}_Q$ são, respectivamente, o vetor posição e o vetor velocidade da i-ésima partícula em relação ao ponto Q. Portanto,

$$\begin{aligned}\frac{d\mathbf{L}_Q}{dt} &= \sum_i m_i \mathbf{v}_i^{(Q)} \times \mathbf{v}_i^{(Q)} + \sum_i (\mathbf{r}_i - \mathbf{r}_Q) \times \dot{\mathbf{p}}_i - \sum_i m_i(\mathbf{r}_i - \mathbf{r}_Q) \times \ddot{\mathbf{r}}_Q \\ &= \sum_{\substack{i,j \\ i\neq j}} (\mathbf{r}_i - \mathbf{r}_Q) \times \mathbf{F}_{ij} + \sum_i (\mathbf{r}_i - \mathbf{r}_Q) \times \mathbf{F}_i^{(e)} - M(\mathbf{R} - \mathbf{r}_Q) \times \ddot{\mathbf{r}}_Q ,\end{aligned} \tag{1.1.11}$$

---

[2] De agora em diante abreviaremos $\sum_{i=1}^N$ por $\sum_i$, estando implícito que a soma estende-se a todas as partículas do sistema, salvo indicação em contrário.

[3] O ponto sobre uma variável denota derivada em relação ao tempo.

tendo sido usadas as Eqs.(1.1.3), (1.1.4) e (1.1.7). Mas, lançando mão novamente da Eq.(A.9), podemos escrever

$$\sum_{\substack{i,j \\ i\neq j}}(\mathbf{r}_i-\mathbf{r}_Q)\times\mathbf{F}_{ij}=\frac{1}{2}\sum_{\substack{i,j \\ i\neq j}}[(\mathbf{r}_i-\mathbf{r}_Q)\times\mathbf{F}_{ij}+(\mathbf{r}_j-\mathbf{r}_Q)\times\mathbf{F}_{ji}]$$

$$=\frac{1}{2}\sum_{\substack{i,j \\ i\neq j}}[(\mathbf{r}_i-\mathbf{r}_Q)-(\mathbf{r}_j-\mathbf{r}_Q)]\times\mathbf{F}_{ij}=\frac{1}{2}\sum_{\substack{i,j \\ i\neq j}}(\mathbf{r}_i-\mathbf{r}_j)\times\mathbf{F}_{ij} \qquad (1.1.12)$$

com a ajuda da Eq.(1.1.2). Se as forças internas obedecem à forma forte da terceira lei de Newton, então $\mathbf{F}_{ij}$ tem a mesma direção que o vetor $\mathbf{r}_{ij}\equiv\mathbf{r}_i-\mathbf{r}_j$ que aponta da j-ésima para a i-ésima partícula, de modo que $\mathbf{r}_{ij}\times\mathbf{F}_{ij}=0$. Assim sendo, a Eq.(1.1.11) pode ser reescrita na forma

$$\frac{d\mathbf{L}_Q}{dt}=\mathbf{N}_Q^{(e)}-M(\mathbf{R}-\mathbf{r}_Q)\times\ddot{\mathbf{r}}_Q\ , \qquad (1.1.13)$$

onde

$$\mathbf{N}_Q^{(e)}=\sum_i(\mathbf{r}_i-\mathbf{r}_Q)\times\mathbf{F}_i^{(e)} \qquad (1.1.14)$$

é o torque externo total em relação ao ponto Q. Se o ponto Q está em repouso ou é o centro de massa, o segundo termo à direita da Eq.(1.1.13) é nulo[4] e ficamos com

$$\frac{d\mathbf{L}_Q}{dt}=\mathbf{N}_Q^{(e)}\ . \qquad (1.1.15)$$

Esta última equação implica uma importante lei de conservação, na qual o momento angular e o torque são tomados relativamente a um ponto fixo ou ao centro de massa.

Teorema da Conservação do Momento Angular. O momento angular total de um sistema de partículas se conserva se o torque externo total é nulo.

A Eq.(1.1.9) assevera que o momento linear total de um sistema de partículas coincide com o calculado como se toda a sua massa estivesse concentrada no centro de massa. No caso do momento angular a situação é um pouco mais complexa. Se $\mathbf{R}$ é o vetor posição do centro de massa em relação à origem O de um sistema de coordenadas inercial e $\mathbf{r}_i'$ é o vetor posição da i-ésima partícula em relação ao centro de massa, então

$$\mathbf{r}_i=\mathbf{r}_i'+\mathbf{R}\quad ,\quad \mathbf{v}_i=\mathbf{v}_i'+\mathbf{V}\ , \qquad (1.1.16)$$

onde $\mathbf{V}\equiv\dot{\mathbf{R}}$ é a velocidade do centro de massa em relação a O e $\mathbf{v}_i'\equiv\dot{\mathbf{r}}_i'$ é a velocidade da i-ésima partícula em relação ao centro de massa. O momento angular total relativamente à origem é

$$\mathbf{L}=\sum_i m_i\mathbf{r}_i\times\mathbf{v}_i=\sum_i m_i\mathbf{r}_i'\times\mathbf{v}_i'+(\sum_i m_i\mathbf{r}_i')\times\mathbf{V}+\mathbf{R}\times\frac{d}{dt}(\sum_i m_i\mathbf{r}_i')+M\mathbf{R}\times\mathbf{V}\ , \qquad (1.1.17)$$

onde usamos (1.1.16). Da primeira das equações (1.1.16) decorre que

[4]Outra possibilidade, que não consideraremos, é a de um ponto Q com aceleração sempre ao longo da linha que une o próprio ponto Q ao centro de massa (Tiersten 1991).

$$\sum_i m_i \mathbf{r}'_i = \sum_i m_i \mathbf{r}_i - \sum_i m_i \mathbf{R} = M\mathbf{R} - M\mathbf{R} = 0 \ . \tag{1.1.18}$$

Assim, o momento angular total em relação à origem admite a decomposição

$$\mathbf{L} = \mathbf{R} \times M\mathbf{V} + \sum_i \mathbf{r}'_i \times \mathbf{p}'_i \ . \tag{1.1.19}$$

Em palavras: o momento angular total em relação à origem é o momento angular do sistema como se estivesse concentrado no centro de massa acrescido do momento angular associado ao movimento em torno do centro de massa.

Por fim, consideremos a energia. A energia cinética total é definida por

$$T = \frac{1}{2} \sum_i m_i v_i^2 \ . \tag{1.1.20}$$

Com o uso de (1.1.16) resulta

$$T = \frac{1}{2} \sum_i m_i v_i'^2 + \frac{1}{2} \sum_i m_i V^2 + \mathbf{V} \cdot \frac{d}{dt} (\sum_i m_i \mathbf{r}'_i) \ , \tag{1.1.21}$$

donde, devido a (1.1.18),

$$T = \frac{M}{2} V^2 + \frac{1}{2} \sum_i m_i v_i'^2 \ . \tag{1.1.22}$$

A energia cinética total é a soma da energia cinética do sistema como se estivesse concentrado no centro de massa com a energia cinética do movimento em torno do centro de massa. Este resultado é particularmente útil na dinâmica do corpo rígido.

O trabalho realizado por todas as forças para levar o sistema de uma configuração inicial A a uma configuração final B é definido por

$$W_{AB} = \sum_i \int_A^B (\mathbf{F}_i^{(e)} + \sum_{\substack{j \\ j \neq i}} \mathbf{F}_{ij}) \cdot d\mathbf{r}_i = \sum_i \int_A^B \mathbf{F}_i^{(e)} \cdot d\mathbf{r}_i + \sum_{\substack{i,j \\ i \neq j}} \int_A^B \mathbf{F}_{ij} \cdot d\mathbf{r}_i \ . \tag{1.1.23}$$

Usando a equação de movimento (1.1.3) deduz-se

$$W_{AB} = \sum_i \int_A^B m_i \dot{\mathbf{v}}_i \cdot \mathbf{v}_i dt = \sum_i \int_A^B d(\frac{1}{2} m_i v_i^2) \ , \tag{1.1.24}$$

donde

$$W_{AB} = T_B - T_A \ , \tag{1.1.25}$$

isto é, o trabalho realizado é igual à variação da energia cinética.

Em numerosos casos as forças são *conservativas*, ou seja, derivam de potenciais escalares. Suponhamos que as forças externas admitam uma função energia potencial $V^{(e)}(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ tal que

$$\mathbf{F}_i^{(e)} = -\boldsymbol{\nabla}_i V^{(e)} \ , \tag{1.1.26}$$

onde $\boldsymbol{\nabla}_i = \hat{\mathbf{i}}\,\partial/\partial x_i + \hat{\mathbf{j}}\,\partial/\partial y_i + \hat{\mathbf{k}}\,\partial/\partial z_i$ é o operador nabla em relação à variável $\mathbf{r}_i$. Neste caso,

$$\sum_i \int_A^B \mathbf{F}_i^{(e)} \cdot d\mathbf{r}_i = -\int_A^B \sum_i \boldsymbol{\nabla}_i V^{(e)} \cdot d\mathbf{r}_i = -\int_A^B dV^{(e)} = V_A^{(e)} - V_B^{(e)} \,. \tag{1.1.27}$$

Se $\mathbf{F}_{ij}$ depende apenas das posições relativas $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ e, além disso, é dedutível de uma função energia potencial $V_{ij}(\mathbf{r}_{ij})$ com $V_{ij} = V_{ji}$, então podemos escrever

$$\mathbf{F}_{ij} = -\boldsymbol{\nabla}_i V_{ij} \,. \tag{1.1.28}$$

Esta forma assegura a validade da versão fraca da lei da ação e reação. Com efeito,

$$\mathbf{F}_{ij} = -\boldsymbol{\nabla}_i V_{ij} = +\boldsymbol{\nabla}_j V_{ij} = +\boldsymbol{\nabla}_j V_{ji} = -\mathbf{F}_{ji} \,. \tag{1.1.29}$$

Se, além disso, V_{ij} depende somente da distância $s_{ij} = |\mathbf{r}_{ij}|$ entre as partículas (forças centrais), temos

$$\mathbf{F}_{ij} = -\boldsymbol{\nabla}_i V_{ij}(s_{ij}) = -\frac{\mathbf{r}_{ij}}{s_{ij}} V'_{ij}(s_{ij}) \tag{1.1.30}$$

onde V'_{ij} é a derivada de V_{ij} em relação a seu argumento, de modo que $\mathbf{F}_{ij}$ aponta ao longo da linha que une as partículas e a lei da ação e reação vale em sua forma forte.

A Eq.(1.1.29) nos permite escrever

$$\sum_{\substack{i,j \\ i\neq j}} \int_A^B \mathbf{F}_{ij} \cdot d\mathbf{r}_i = \frac{1}{2} \sum_{\substack{i,j \\ i\neq j}} \int_A^B (\mathbf{F}_{ij} \cdot d\mathbf{r}_i + \mathbf{F}_{ji} \cdot d\mathbf{r}_j) = \frac{1}{2} \sum_{\substack{i,j \\ i\neq j}} \int_A^B \mathbf{F}_{ij} \cdot d(\mathbf{r}_i - \mathbf{r}_j)$$

$$= -\frac{1}{2} \sum_{\substack{i,j \\ i\neq j}} \int_A^B \boldsymbol{\nabla}_i V_{ij} \cdot d\mathbf{r}_{ij} = -\frac{1}{2} \sum_{\substack{i,j \\ i\neq j}} \int_A^B \boldsymbol{\nabla}_{ij} V_{ij} \cdot d\mathbf{r}_{ij} = -\frac{1}{2} \sum_{\substack{i,j \\ i\neq j}} V_{ij} \Big|_A^B \,, \tag{1.1.31}$$

onde $\boldsymbol{\nabla}_{ij}$ denota o gradiente em relação ao vetor $\mathbf{r}_{ij}$, tendo sido utilizada a propriedade evidente $\boldsymbol{\nabla}_i V_{ij} = \boldsymbol{\nabla}_{ij} V_{ij}$. Finalmente, combinando as Eqs.(1.1.23), (1.1.25), (1.1.27) e (1.1.31) deduz-se

$$(T + V)_A = (T + V)_B \,, \tag{1.1.32}$$

com

$$V = V^{(e)} + \frac{1}{2} \sum_{\substack{i,j \\ i\neq j}} V_{ij} \,. \tag{1.1.33}$$

O segundo termo no lado direito da Eq.(1.1.33) é chamado de energia potencial interna do sistema. Destes últimos resultados decorre uma importante lei de conservação.[5]

Teorema da Conservação da Energia. Se todas as forças são conservativas, a energia total $E = T + V$ de um sistema de partículas é conservada.

[5] É evidente que os teoremas de conservação desta seção valem somente em referenciais inerciais, embora isto não conste explicitamente nos referidos enunciados.

1.2 Vínculos

Vínculos são limitações às possíveis posições e velocidades das partículas de um sistema mecânico, restringindo *a priori* o seu movimento. É importante sublinhar que vínculos são limitações *de ordem cinemática* impostas ao sistema mecânico. Tais restrições, portanto, antecedem a dinâmica e precisam ser levadas em conta na formulação das equações de movimento do sistema. Restrições de natureza dinâmica – decorrentes, portanto, das equações de movimento – *não* são vínculos. Por exemplo, a segunda lei de Newton obriga uma partícula sujeita a uma força central a se mover num plano fixo, mas isto não caracteriza um vínculo.

Os exemplos considerados a seguir são bastante simples, mas ilustram alguns dos aspectos essenciais dos vínculos.

■ **Exemplo 1.2.1.** Uma partícula está restrita a uma superfície fixa. Seja $\mathbf{r} = (x, y, z)$ o vetor posição da partícula relativamente a um sistema cartesiano de eixos em relação ao qual a superfície permanece fixa. Então x, y, z não são variáveis independentes mas devem satisfazer

$$f(\mathbf{r}) \equiv f(x, y, z) = 0 \ , \tag{1.2.1}$$

onde $f(\mathbf{r}) = 0$ é a equação da superfície. Se, por exemplo, a superfície for uma esfera centrada na origem,

$$f(x, y, z) = x^2 + y^2 + z^2 - R^2 \ , \tag{1.2.2}$$

onde R é o raio da esfera. ■

■ **Exemplo 1.2.2.** Uma partícula está restrita a uma superfície móvel ou deformável. Neste caso x, y, z obedecem à equação

$$f(\mathbf{r}, t) \equiv f(x, y, z, t) = 0 \ , \tag{1.2.3}$$

a dependência temporal explícita indicando a mudança da forma ou localização da superfície no transcurso do tempo. ■

■ **Exemplo 1.2.3.** Duas partículas movem-se no espaço sempre unidas por uma haste rígida. O vínculo tem a forma

$$(\mathbf{r}_2 - \mathbf{r}_1)^2 - l^2 = 0 \tag{1.2.4}$$

ou, equivalentemente,

$$(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 - l^2 = 0 \ , \tag{1.2.5}$$

sendo l o comprimento invariável da haste. ■

■ **Exemplo 1.2.4.** Um pêndulo duplo oscila num plano vertical fixo. As equações de vínculo são

$$x_1^2 + y_1^2 - l_1^2 = 0 \ , \ (x_2 - x_1)^2 + (y_2 - y_1)^2 - l_2^2 = 0 \ , \tag{1.2.6}$$

como se depreende da Fig. 1.1. ■

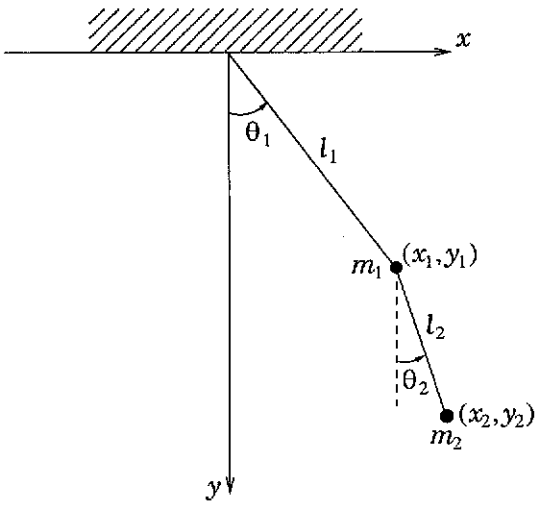


Figura 1.1: Pêndulo duplo plano.

Vínculos Holônomos

Todos os vínculos discutidos acima são holônomos. Se $\xi_1, \ldots, \xi_M$ são coordenadas arbitrárias usadas para descrever a *configuração*[6] de um sistema mecânico, um vínculo é chamado de *holônomo*[7] quando pode ser expresso por uma equação da forma

$$f(\xi_1, \ldots, \xi_M, t) = 0 \,, \tag{1.2.7}$$

isto é, por uma relação funcional exclusivamente entre as coordenadas, podendo envolver o tempo de modo explícito. Vínculos que não podem ser assim representados são ditos não-holônomos. Por exemplo, a imposição de que as moléculas de um gás permaneçam no interior de um recipiente é descrita por desigualdades: se o recipiente é uma caixa de arestas a, b, c temos $0 < x_i < a$, $0 < y_i < b$, $0 < z_i < c$, onde $\mathbf{r}_i = (x_i, y_i, z_i)$ é o vetor posição da i-ésima molécula (tratada como partícula puntiforme). Ocorrem com freqüência, especialmente na dinâmica do corpo rígido, vínculos representáveis por equações envolvendo velocidades, isto é, equações diferenciais da forma

$$g(\xi_1, \ldots, \xi_M, \dot{\xi}_1, \ldots, \dot{\xi}_M, t) = 0 \,. \tag{1.2.8}$$

■ **Exemplo 1.2.5.** Um cilindro rola sem deslizar ao longo de uma linha reta. Sendo x a posição do centro de massa do cilindro e ϕ o ângulo de rotação em torno do centro de massa, a condição de rolar sem deslizar representa-se por

$$\dot{x} = R\dot{\phi} \,, \tag{1.2.9}$$

[6] A posição de cada uma das partículas de um sistema mecânico num dado instante define a configuração do sistema no referido instante.

[7] Do grego *hólos* (inteiro, completo) e *nómos* (regra, lei).

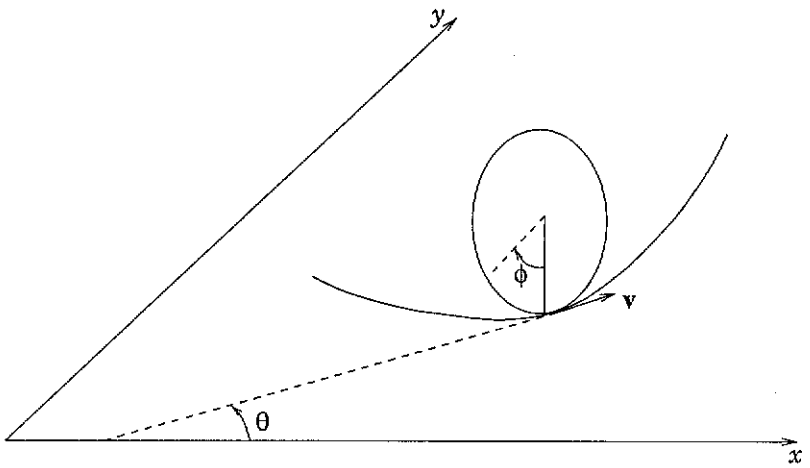


Figura 1.2: Disco rolando sem deslizar num plano horizontal.

onde R é o raio do cilindro. ■

■ **Exemplo 1.2.6.** Um disco vertical rola sem deslizar num plano horizontal. Sejam (x, y) a posição do centro do disco, θ o ângulo do plano do disco com o eixo x e ϕ o ângulo de rotação do disco em torno do seu eixo de simetria (ver Fig. 1.2). Sendo v a velocidade do centro de massa, o disco rola sem deslizar desde que $v = R\dot{\phi}$. Notando que $\dot{x} \equiv v_x = v\cos\theta$ e $\dot{y} \equiv v_y = v\,\mathrm{sen}\,\theta$, somos conduzidos às equações

$$\dot{x} - R\dot{\phi}\cos\theta = 0 \quad , \quad \dot{y} - R\dot{\phi}\,\mathrm{sen}\,\theta = 0 \ , \tag{1.2.10}$$

que exprimem matematicamente a condição de rolamento sem deslizamento. ■

As equações (1.2.9) e (1.2.10) são exatamente da forma (1.2.8). Em geral, vínculos deste tipo não podem ser reduzidos por uma integração à forma (1.2.7) e, conseqüentemente, não são holônomos. Vínculos não-integráveis da forma (1.2.8) restringem os deslocamentos possíveis do sistema mas não impõem quaisquer limites às configurações possíveis, de modo que *a priori* todas as configurações são acessíveis. Se um vínculo originalmente expresso na forma (1.2.8) puder ser reduzido por uma integração à forma (1.2.7), ele será dito holônomo. Por exemplo, o vínculo (1.2.9) é holônomo porque equivale a $x - R\phi = 0$, que tem a forma (1.2.7).[8] No entanto, vínculos diferenciais do tipo (1.2.8) raramente são integráveis, como o Exemplo 1.2.6 ilustra bem.

■ **Exemplo 1.2.7.** Mostrar que cada um dos vínculos (1.2.10) não é integrável.

Solução. A primeira das Eqs.(1.2.10) é integrável se existe um *fator integrante* não-nulo $h(x, y, \theta, \phi, t)$ tal que

[8] A constante de integração pode ser feita igual a zero por uma escolha adequada da posição ou do ângulo no instante inicial.

$$h\dot{x} - hR\dot{\phi}\cos\theta = \frac{d}{dt}G(x,y,\theta,\phi,t) = \frac{\partial G}{\partial x}\dot{x} + \frac{\partial G}{\partial y}\dot{y} + \frac{\partial G}{\partial \theta}\dot{\theta} + \frac{\partial G}{\partial \phi}\dot{\phi} + \frac{\partial G}{\partial t} \tag{1.2.11}$$

para alguma função G, pois neste caso, em virtude de (1.2.10), a Eq.(1.2.11) será equivalente a (C é uma constante arbitrária)

$$G(x,y,\theta,\phi,t) - C = 0\,, \tag{1.2.12}$$

que tem a forma (1.2.7). Mas, admitindo que o primeiro membro de (1.2.11) seja uma derivada total, é necessário levar em conta que a ordem de diferenciação é irrelevante nas derivadas segundas cruzadas da função supostamente existente G. De acordo com a hipótese (1.2.11) teríamos $\partial G/\partial x = h$ e $\partial G/\partial\phi = -hR\cos\theta$, as demais derivadas parciais sendo nulas. Logo,

$$\frac{\partial^2 G}{\partial y \partial x} = \frac{\partial^2 G}{\partial x \partial y} \implies \frac{\partial h}{\partial y} = 0 \implies h = h(x,\theta,\phi,t)\,;$$

$$\frac{\partial^2 G}{\partial \theta \partial x} = \frac{\partial^2 G}{\partial x \partial \theta} \implies \frac{\partial h}{\partial \theta} = 0 \implies h = h(x,\phi,t)\,;$$

$$\frac{\partial^2 G}{\partial \theta \partial \phi} = \frac{\partial^2 G}{\partial \phi \partial \theta} \implies \frac{\partial}{\partial \theta}(-hR\cos\theta) = 0 \implies hR\,\mathrm{sen}\,\theta = 0 \implies h = 0\,.$$

A inexistência de fator integrante não-nulo prova que a primeira das Eqs.(1.2.10) não constitui um vínculo holônomo, argumento análogo aplicando-se à segunda das Eqs.(1.2.10). ■

Vínculos expressos por equações diferenciais não-integráveis representam restrições sobre as velocidades. O centro de massa do disco do Exemplo 1.2.6 pode percorrer uma circunferência de tal modo que x, y e θ retornem aos seus valores originais. Mas, ao final do processo, ϕ pode atingir um valor arbitrário, dependendo do comprimento da circunferência descrita. Isto indica que não existe qualquer relação de dependência funcional entre as variáveis que descrevem a configuração do disco em cada instante.

1.3 Princípio de d'Alembert

A configuração de um sistema mecânico é definida pela posição de cada uma das partículas do sistema num determinado momento. Para um sistema mecânico sujeito a vínculos, num dado instante t há uma infinidade de configurações possíveis, isto é, consistentes com os vínculos.

Deslocamentos Virtuais

Os deslocamentos infinitesimais de cada partícula que levam de uma configuração possível a outra configuração possível infinitesimalmente próxima no *mesmo instante* t são chamados de *deslocamentos virtuais*. Mais precisamente, dado um sistema de N partículas os deslocamentos virtuais $\delta\mathbf{r}_i$, $i = 1,\ldots,N$, são deslocamentos infinitesimais das posições $\mathbf{r}_1,\ldots,\mathbf{r}_N$ realizados instantaneamente[9] e com a propriedade de serem compatíveis com os vínculos. Em suma, as características definidoras dos deslocamentos virtuais são: (i) eles são infinitesimais; (ii) ocorrem num instante t fixo; (iii) não violam os vínculos.

[9] A exigência de que o tempo permaneça fixo confere um caráter fictício aos deslocamentos virtuais, embora em muitas situações eles possam corresponder a deslocamentos verdadeiros que ocorrem num intervalo de tempo dt, conforme se verá nos exemplos subseqüentes. Nos casos em que todos os vínculos são independentes do tempo, os deslocamentos virtuais coincidem com os deslocamentos reais.

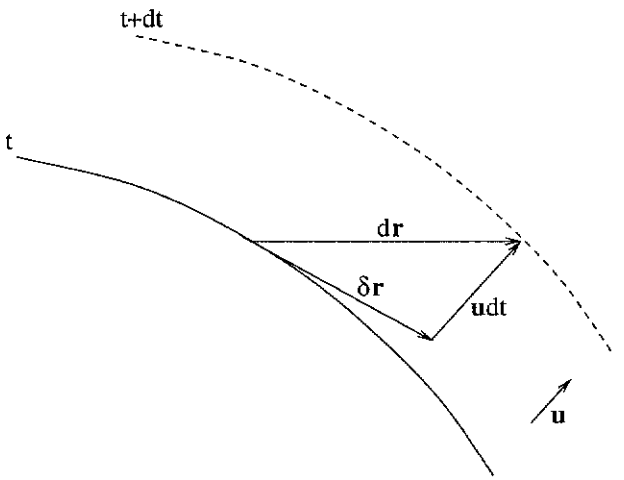


Figura 1.3: Deslocamentos virtual e real de partícula em superfície móvel.

■ **Exemplo 1.3.1.** Um partícula está restrita a uma superfície móvel. Seja $f(\mathbf{r},t)=0$ a equação da superfície. Um deslocamento virtual deve ser consistente com o vínculo, isto é, o ponto $\mathbf{r}$ e o ponto deslocado $\mathbf{r}+\delta\mathbf{r}$ devem pertencer à superfície no mesmo instante t:

$$f(\mathbf{r}+\delta\mathbf{r},t)=0 \implies f(\mathbf{r},t)+\nabla f\cdot\delta\mathbf{r}=0 \implies \nabla f\cdot\delta\mathbf{r}=0\,. \tag{1.3.1}$$

Como ∇f é perpendicular à superfície no instante t, o deslocamento virtual $\delta\mathbf{r}$ é tangente à superfície nesse instante (Fig. 1.3). ■

Note, contudo, que um deslocamento real $d\mathbf{r}$ se dá num intervalo de tempo dt. Portanto, para que a partícula permaneça na superfície é preciso que

$$f(\mathbf{r}+d\mathbf{r},t+dt)=0 \implies \nabla f\cdot d\mathbf{r}+\frac{\partial f}{\partial t}\,dt=0\,. \tag{1.3.2}$$

Vê-se, assim, que $d\mathbf{r}$ não é tangente à superfície se $\partial f/\partial t\neq 0$. Somente o deslocamento virtual realizado a tempo fixo é tangente à superfície mesmo que ela esteja em movimento. A Fig. 1.3 ilustra a diferença entre deslocamentos virtual e real para uma partícula restrita a uma superfície cuja velocidade no instante t é $\mathbf{u}$.

Trabalho Virtual

A importância de se introduzir a noção de deslocamento virtual advém da seguinte observação simples. Se a superfície à qual o movimento da partícula está restrito é idealmente lisa, a força de contato entre a partícula e a superfície não possui componente tangencial, mas apenas normal à superfície. Desta forma, o trabalho realizado pela força de vínculo por ocasião de um um deslocamento virtual da partícula é nulo mesmo que a superfície esteja em movimento, diferentemente do trabalho realizado durante um deslocamento real, que não é necessariamente zero. Na maioria dos casos fisicamente interessantes o trabalho virtual *total* das forças de vínculo se anula, como os próximos exemplos atestam.

■ **Exemplo 1.3.2.** Duas partículas movem-se unidas por uma haste rígida. Sejam $\mathbf{f}_1$ e $\mathbf{f}_2$ as forças de vínculo sobre as partículas, com $\mathbf{f}_1 = -\mathbf{f}_2$, pela terceira lei de Newton, e tanto $\mathbf{f}_1$ quanto $\mathbf{f}_2$ dirigidas ao longo da reta definida pelas partículas. O trabalho virtual das forças de vínculo é

$$\delta W_v = \mathbf{f}_1 \cdot \delta\mathbf{r}_1 + \mathbf{f}_2 \cdot \delta\mathbf{r}_2 = \mathbf{f}_2 \cdot (\delta\mathbf{r}_2 - \delta\mathbf{r}_1) \ .$$

Definindo $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$, a equação de vínculo assume a forma (1.2.4), a saber, $r^2 - l^2 = 0$. Em termos da variável $\mathbf{r}$ a situação equivale àquela discutida no Exemplo 1.3.1. Tomando $f(\mathbf{r}, t) = r^2 - l^2$ a Eq.(1.3.1) reduz-se a $\mathbf{r} \cdot \delta\mathbf{r} = 0$. Mas, como $\mathbf{f}_2$ e $\mathbf{r}$ são colineares, existe um escalar λ tal que $\mathbf{f}_2 = \lambda\mathbf{r}$, donde $\delta W_v = \lambda\mathbf{r} \cdot \delta\mathbf{r} = 0$. Uma vez que um corpo rígido consiste num vasto número de partículas cujas distâncias mútuas são invariáveis, conclui-se que é zero o trabalho virtual das forças responsáveis pela rigidez do corpo. ■

■ **Exemplo 1.3.3.** Um corpo rígido rola sem deslizar sobre uma superfície fixa. De modo geral, para que não haja deslizamento é preciso que exista uma força de atrito entre o corpo e a superfície, isto é, as superfícies em contato devem ser ásperas. Mas, ao rolar sem deslizar, em cada instante as partículas do corpo giram em torno de um eixo que passa pelo ponto de contato do corpo com a superfície. Assim, a força de atrito atua sobre um ponto do corpo que em cada instante possui velocidade nula, pois encontra-se sobre o eixo instantâneo de rotação. Uma vez que qualquer deslocamento virtual do corpo não pode mover o seu ponto de contato com a superfície, senão haveria deslizamento com a conseqüente violação do vínculo, o trabalho virtual da força de vínculo é $\delta W_v = \mathbf{f} \cdot \delta\mathbf{r} = 0$ porque $\delta\mathbf{r} = 0$, muito embora $\mathbf{f} \neq 0$. ■

A análise anterior torna patente que numa ampla gama de situações fisicamente relevantes o trabalho virtual total das forças de vínculo é nulo. A situação em que há forças de atrito de deslizamento constitui uma exceção. Neste caso, a força de vínculo possui componente na mesma direção do deslocamento virtual e o trabalho virtual por ela realizado não é zero. A noção de atrito é estritamente macroscópica, portanto de escasso interesse para o desenvolvimento das formulações gerais da mecânica, especialmente sob a óptica da física contemporânea. Vínculos cujas forças associadas não realizam trabalho durante deslocamentos virtuais são chamados de *vínculos ideais*. De agora em diante vamos nos limitar, sem perda significativa de generalidade, à consideração de sistemas mecânicos sujeitos apenas a vínculos ideais.

Princípio dos Trabalhos Virtuais

A formulação newtoniana da mecânica caracteriza-se pelo conjunto de equações diferenciais

$$m_i\ddot{\mathbf{r}}_i = \mathbf{F}_i \quad , \quad i = 1, \ldots, N \quad , \tag{1.3.3}$$

onde $\mathbf{F}_i$ é a força *total* ou *resultante* sobre a i-ésima partícula, supostamente uma função conhecida das posições, velocidades e tempo. Este sistema de equações diferenciais determina uma única solução para os $\mathbf{r}_i(t)$ uma vez especificadas todas as posições e velocidades num dado instante inicial.

Quando há vínculos em ação, saltam aos olhos os inconvenientes da formulação newtoniana. Em primeiro lugar, ela geralmente requer o uso de mais coordenadas do que o necessário para especificar a configuração do sistema em cada instante. Quando há vínculos holônomos, por exemplo, as posições $\mathbf{r}_1, \ldots, \mathbf{r}_N$ não são independentes entre

si, tornando a formulação newtoniana antieconômica ao exigir o emprego de variáveis redundantes. Além disso, a força total sobre a i-ésima partícula admite a decomposição

$$\mathbf{F}_i = \mathbf{F}_i^{(a)} + \mathbf{f}_i \ , \tag{1.3.4}$$

onde $\mathbf{F}_i^{(a)}$ é a força *aplicada* e $\mathbf{f}_i$ é a força de vínculo. No caso do pêndulo duplo do Exemplo 1.2.4 , $\mathbf{F}_1^{(a)}$ e $\mathbf{F}_2^{(a)}$ são os pesos das partículas, ao passo que $\mathbf{f}_1$ e $\mathbf{f}_2$ são determinadas pelas tensões nos fios. A dificuldade, aqui, reside em desconhecer-se *a priori* as expressões das forças de vínculo como funções das posições e velocidades, pois o que se conhece, na verdade, são os *efeitos* produzidos pelas forças de vínculo. Pode-se argumentar, ainda, que as forças aplicadas é que costumam ser as verdadeiras causas do movimento, as forças de vínculo servindo meramente para assegurar a preservação das restrições geométricas ou cinemáticas no decurso do tempo. Não menos importante é o fato de as leis de Newton revelarem-se incapazes de descrever corretamente o movimento de certos sistemas sujeitos a vínculos (Stadler 1982).

Por todos esses motivos, seria desejável obter uma formulação da mecânica clássica que fosse a mais parcimoniosa possível, a saber, envolvesse somente forças aplicadas e empregasse apenas coordenadas mutuamente independentes. Veremos que tal objetivo é alcançado pelo formalismo lagrangiano quando todos os vínculos são holônomos. Como passo intermediário para chegar à formulação de Lagrange, iremos discutir o chamado princípio de d'Alembert, que é um método de escrever as equações de movimento exclusivamente em termos das forças aplicadas, e para cuja dedução será explorado o fato de que o trabalho virtual das forças de vínculo é nulo.

Consideremos, inicialmente, uma situação estática, isto é, um sistema de partículas em equilíbrio. Neste caso $\mathbf{F}_i = 0$ e, quaisquer que sejam os deslocamentos virtuais $\delta\mathbf{r}_i$,

$$\sum_i \mathbf{F}_i \cdot \delta\mathbf{r}_i = 0 \ . \tag{1.3.5}$$

Fazendo uso da decomposição (1.3.4) resulta

$$\sum_i \mathbf{F}_i^{(a)} \cdot \delta\mathbf{r}_i + \sum_i \mathbf{f}_i \cdot \delta\mathbf{r}_i = 0 \ . \tag{1.3.6}$$

Limitando-nos ao conjunto suficientemente amplo de circunstâncias em que o trabalho virtual das forças de vínculo é zero, somos conduzidos ao chamado *princípio dos trabalhos virtuais*:

$$\sum_i \mathbf{F}_i^{(a)} \cdot \delta\mathbf{r}_i = 0 \ . \tag{1.3.7}$$

Este princípio permite exprimir a condição de equilíbrio para sistemas vinculados em termos somente das forças aplicadas.[10]

Princípio de d'Alembert

Estamos interessados na dinâmica, que pode ser formalmente reduzida à estática escrevendo a segunda lei de Newton na forma $\mathbf{F}_i - \dot{\mathbf{p}}_i = 0$, com $\mathbf{p}_i = m_i\dot{\mathbf{r}}_i$. Segundo a

[10] Formulado por Johann Bernoulli em 1717, o princípio dos trabalhos virtuais remonta a Jordanus, no século XIII (Dugas 1988). Ilustrações do seu emprego em casos interessantes encontram-se em Sommerfeld (1952) e Synge & Griffith (1959).

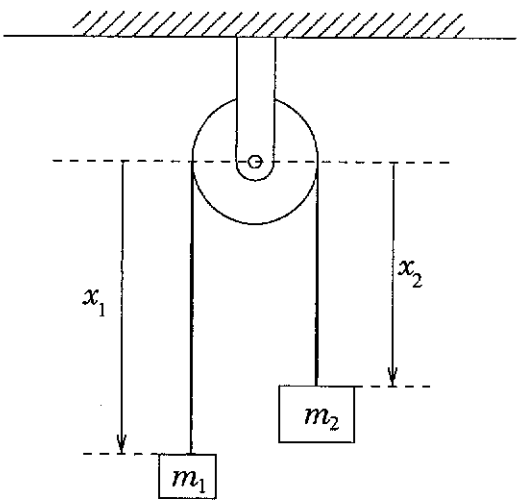


Figura 1.4: Máquina de Atwood.

interpretação de d'Alembert, cada partícula do sistema encontra-se em "equilíbrio" sob uma força resultante que é a soma da força real com uma "força efetiva invertida" igual a $-\dot{\mathbf{p}}_i$. Esta força adicional fictícia é uma força de inércia existente no referencial que acompanha o movimento da partícula, isto é, no qual ela permanece em repouso (Lanczos 1970; Sommerfeld 1952). Interpretações à parte, o fato é que agora, em lugar de (1.3.5), a equação

$$\sum_i (\dot{\mathbf{p}}_i - \mathbf{F}_i) \cdot \delta \mathbf{r}_i = 0 \tag{1.3.8}$$

é obviamente verdadeira quaisquer que sejam os deslocamentos virtuais $\delta \mathbf{r}_i$. Usando novamente a decomposição (1.3.4) e admitindo a nulidade do trabalho virtual das forças de vínculo, resulta o chamado *princípio de d'Alembert*:

$$\sum_i (\dot{\mathbf{p}}_i - \mathbf{F}_i^{(a)}) \cdot \delta \mathbf{r}_i = 0 \ . \tag{1.3.9}$$

Este princípio representa uma extensão do princípio dos trabalhos virtuais a sistemas mecânicos em movimento. No caso de sistemas vinculados, o princípio de d'Alembert constitui um avanço relativamente à formulação newtoniana porque exclui qualquer referência às forças de vínculo. Em suas aplicações concretas, no entanto, é preciso levar em conta que os deslocamentos virtuais $\delta \mathbf{r}_i$ não são independentes, pois têm que estar em harmonia com os vínculos.

■ **Exemplo 1.3.4.** Utilizando o princípio de d'Alembert, encontrar as equações de movimento para o sistema mecânico da Fig. 1.4, conhecido como máquina de Atwood.

Solução. A roldana da Fig. 1.4 é suposta sem massa e sem atrito com o eixo. Com o sistema cartesiano indicado na figura, temos $\mathbf{r}_1 = x_1 \hat{\mathbf{i}}$, $\mathbf{r}_2 = x_2 \hat{\mathbf{i}}$ e o vínculo holônomo escreve-se

$$x_1 + x_2 = l \ , \tag{1.3.10}$$

onde a constante l é determinada pelo raio da roldana e pelo comprimento do fio, suposto inextensível e de massa desprezível. Claramente, os deslocamentos virtuais δx_1 e δx_2 compatíveis com o vínculo (1.3.10) estão relacionados por

$$\delta x_1 + \delta x_2 = 0 \implies \delta x_2 = -\delta x_1 \ .$$

Em outras palavras, se uma das massas sobe a outra desce a mesma distância, e vice-versa. Em virtude destas últimas equações, temos $\delta \mathbf{r}_1 = \delta x_1 \hat{\mathbf{i}}$ e $\delta \mathbf{r}_2 = \delta x_2 \hat{\mathbf{i}} = -\delta x_1 \hat{\mathbf{i}} = -\delta \mathbf{r}_1$. Notando que $\ddot{\mathbf{r}}_1 = \ddot{x}_1 \hat{\mathbf{i}}$, $\ddot{\mathbf{r}}_2 = \ddot{x}_2 \hat{\mathbf{i}}$ e levando em conta, ainda, que $\ddot{x}_2 = -\ddot{x}_1$ (como se deduz derivando duas vezes a Eq.(1.3.10) em relação ao tempo), o princípio de d'Alembert

$$m_1 \ddot{\mathbf{r}}_1 \cdot \delta \mathbf{r}_1 + m_2 \ddot{\mathbf{r}}_2 \cdot \delta \mathbf{r}_2 = \mathbf{F}_1^{(a)} \cdot \delta \mathbf{r}_1 + \mathbf{F}_2^{(a)} \cdot \delta \mathbf{r}_2 = m_1 g \hat{\mathbf{i}} \cdot \delta \mathbf{r}_1 + m_2 g \hat{\mathbf{i}} \cdot \delta \mathbf{r}_2$$

reduz-se a

$$m_1 \ddot{x}_1 \delta x_1 + (-m_2 \ddot{x}_1)(-\delta x_1) = m_1 g \delta x_1 + m_2 g (-\delta x_1) \implies (m_1 + m_2) \ddot{x}_1 \delta x_1 = (m_1 - m_2) g \delta x_1 \ .$$

Em vista da arbitrariedade de δx_1, resulta a equação de movimento da massa m_1:

$$(m_1 + m_2) \ddot{x}_1 = (m_1 - m_2) g \ .$$

A aceleração da massa m_1 é

$$\ddot{x}_1 = \frac{m_1 - m_2}{m_1 + m_2} g \ ,$$

que coincide com o resultado obtido pelo tratamento newtoniano elementar. A aceleração de m_2 é simplesmente $\ddot{x}_2 = -\ddot{x}_1$. ■

Para outras aplicações do princípio de d'Alembert o leitor é remetido a Sommerfeld (1952) e Synge & Griffith (1959).

1.4 Coordenadas Generalizadas e Equações de Lagrange

O princípio de d'Alembert ainda exige trabalhar com mais coordenadas do que o necessário, pois não apenas os $\delta \mathbf{r}_i$ não são independentes como também, em muitos casos, as próprias posições $\mathbf{r}_i$.

Coordenadas Generalizadas

Em sistemas holônomos é possível introduzir um certo número n de variáveis independentes, denotadas genericamente por $q_1, \ldots, q_n$ e denominadas *coordenadas generalizadas*, de sorte que: (a) o vetor posição de cada partícula é determinado univocamente em cada instante pelos valores dos qs; (b) os vínculos, supostos todos da forma (1.2.7), são *identicamente* satisfeitos se expressos em termos dos qs. Vejamos dois casos ilustrativos.

■ **Exemplo 1.4.1.** No caso do pêndulo duplo plano, discutido no Exemplo 1.2.4, uma escolha possível de coordenadas generalizadas é $q_1 = \theta_1$, $q_2 = \theta_2$ (vide Fig. 1.1). Temos, então,

$$x_1 = l_1 \operatorname{sen} \theta_1 \ , \ y_1 = l_1 \cos \theta_1 \ , \ x_2 = l_1 \operatorname{sen} \theta_1 + l_2 \operatorname{sen} \theta_2 \ , \ y_2 = l_1 \cos \theta_1 + l_2 \cos \theta_2 \ .$$

Note que os valores de θ_1 e θ_2 especificam univocamente as posições das partículas, isto é, a configuração do sistema. Em termos de θ_1 e θ_2 as equações de vínculo (1.2.6) reduzem-se às identidades $l_1^2 \operatorname{sen}^2\theta_1 + l_1^2 \cos^2\theta_1 - l_1^2 = 0$ e $l_2^2 \operatorname{sen}^2\theta_2 + l_2^2 \cos^2\theta_2 - l_2^2 = 0$. ■

■ **Exemplo 1.4.2.** Uma partícula está obrigada a permanecer sobre a superfície de uma esfera em movimento uniforme. Seja $\mathbf{u} = (u_x, u_y, u_z)$ a velocidade constante da esfera em relação a um referencial inercial. No instante t o centro da esfera tem coordenadas $(u_x t, u_y t, u_z t)$ e a equação de vínculo tem a forma

$$(x - u_x t)^2 + (y - u_y t)^2 + (z - u_z t)^2 - R^2 = 0 \ ,$$

sendo R o raio da esfera. Introduzindo os ângulos θ e φ pelas equações

$$x = u_x t + R \operatorname{sen}\theta \cos\varphi \ , \ y = u_y t + R \operatorname{sen}\theta \operatorname{sen}\varphi \ , \ z = u_z t + R\cos\theta \ ,$$

a equação de vínculo passa a ser identicamente satisfeita. Portanto, $q_1 = \theta$ e $q_2 = \varphi$ constitui uma escolha possível de coordenadas generalizadas. ■

Seja um sistema mecânico constituído por N partículas submetidas aos p vínculos holônomos

$$\begin{aligned} f_1(\mathbf{r}_1, \ldots, \mathbf{r}_N, t) &= 0 \ , \\ &\vdots \\ f_p(\mathbf{r}_1, \ldots, \mathbf{r}_N, t) &= 0 \ . \end{aligned} \tag{1.4.1}$$

Das $3N$ coordenadas $(x_1, y_1, z_1), \ldots, (x_N, y_N, z_N)$ apenas $n = 3N - p$ podem ser tomadas como independentes entre si, e diz-se que o sistema possui n *graus de liberdade*. É possível introduzir n coordenadas generalizadas $q_1, \ldots, q_n$ em termos das quais

$$\mathbf{r}_i = \mathbf{r}_i(q_1, \ldots, q_n, t) \ , \ i = 1, \ldots, N \ , \tag{1.4.2}$$

e as Eqs.(1.4.1) são identicamente satisfeitas. Em linguagem geométrica, pode-se dizer que as Eqs.(1.4.1) definem uma hipersuperfície de dimensão n num espaço de dimensão $3N$ e que (1.4.2) são as equações paramétricas dessa hipersuperfície.

Cada conjunto de valores atribuídos às coordenadas generalizadas define uma configuração do sistema, isto é, as posições de todas as partículas num dado instante. O espaço cartesiano que tem as coordenadas generalizadas como eixos coordenados é chamado de *espaço de configuração* do sistema. A representação do espaço de configuração como um espaço cartesiano é apenas simbólica, no entanto. Por exemplo, no caso de uma partícula restrita à superfície de uma esfera e descrita pelas coordenadas esféricas angulares (θ, φ), uma única configuração corresponde à infinidade de pontos $\theta = 0$ com φ arbitrário. A rigor, o espaço de configuração possui a estrutura matemática de uma variedade diferenciável, daí ser também chamado de *variedade de configuração*. Um sistema mecânico holônomo tem tantos graus de liberdade quantas sejam as coordenadas generalizadas necessárias e suficientes para especificar a sua configuração em cada instante.

Uma vez introduzidas coordenadas generalizadas via (1.4.2), os deslocamentos virtuais $\delta\mathbf{r}_i$ podem ser expressos em termos dos deslocamentos virtuais independentes δq_k mediante

$$\delta\mathbf{r}_i = \sum_{k=1}^{n} \frac{\partial \mathbf{r}_i}{\partial q_k} \delta q_k \ , \tag{1.4.3}$$

já que o tempo deve permanecer fixo.[11] Por outro lado,

$$\mathbf{v}_i = \frac{d\mathbf{r}_i}{dt} = \sum_{k=1}^{n} \frac{\partial \mathbf{r}_i}{\partial q_k}\dot{q}_k + \frac{\partial \mathbf{r}_i}{\partial t} \ . \tag{1.4.4}$$

Tendo em vista que as forças de vínculo não aparecem no princípio de d'Alembert, de ora em diante abandonaremos o superescrito identificador das forças aplicadas, isto é, adotaremos a notação abreviada $\mathbf{F}_i^{(a)} \equiv \mathbf{F}_i$. Usando (1.4.3), o trabalho virtual das forças aplicadas torna-se

$$\sum_{i=1}^{N} \mathbf{F}_i \cdot \delta\mathbf{r}_i = \sum_{i=1}^{N}\sum_{k=1}^{n} \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_k}\delta q_k \equiv \sum_{k=1}^{n} Q_k\, \delta q_k \ , \tag{1.4.5}$$

onde

$$Q_k = \sum_{i=1}^{N} \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_k} \tag{1.4.6}$$

é, por definição, a k-ésima componente da *força generalizada*. Como os qs não têm necessariamente dimensão de comprimento, os Qs não têm necessariamente dimensão de força; mas cada termo $Q_k\,\delta q_k$ tem sempre dimensão de trabalho.

A outra quantidade envolvida no princípio de d'Alembert é

$$\sum_{i=1}^{N} \dot{\mathbf{p}}_i \cdot \delta\mathbf{r}_i = \sum_{i=1}^{N} m_i\dot{\mathbf{v}}_i \cdot \delta\mathbf{r}_i = \sum_{i=1}^{N}\sum_{k=1}^{n} m_i\dot{\mathbf{v}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_k}\delta q_k \ . \tag{1.4.7}$$

A seguinte identidade será útil:

$$\sum_{i=1}^{N} m_i\dot{\mathbf{v}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_k} = \sum_{i=1}^{N}\left\{ \frac{d}{dt}\left(m_i\mathbf{v}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_k}\right) - m_i\mathbf{v}_i \cdot \frac{d}{dt}\left(\frac{\partial \mathbf{r}_i}{\partial q_k}\right)\right\} \ . \tag{1.4.8}$$

No último termo de (1.4.8) podemos usar o resultado

$$\frac{d}{dt}\left(\frac{\partial \mathbf{r}_i}{\partial q_k}\right) = \sum_{l=1}^{n} \frac{\partial}{\partial q_l}\left(\frac{\partial \mathbf{r}_i}{\partial q_k}\right)\dot{q}_l + \frac{\partial}{\partial t}\left(\frac{\partial \mathbf{r}_i}{\partial q_k}\right) = \frac{\partial}{\partial q_k}\left(\sum_{l=1}^{n} \frac{\partial \mathbf{r}_i}{\partial q_l}\dot{q}_l + \frac{\partial \mathbf{r}_i}{\partial t}\right) = \frac{\partial \mathbf{v}_i}{\partial q_k} \ , \tag{1.4.9}$$

onde utilizamos (1.4.4) e passamos a tratar os qs e $\dot{q}$s como grandezas independentes, de modo que as derivadas parciais em relação aos qs tratam os $\dot{q}$s como constantes e vice-versa. Além disso, de (1.4.4) deduz-de imediatamente

$$\frac{\partial \mathbf{v}_i}{\partial \dot{q}_k} = \frac{\partial \mathbf{r}_i}{\partial q_k} \ , \tag{1.4.10}$$

permitindo escrever (1.4.8) na forma

$$\sum_{i=1}^{N} m_i\dot{\mathbf{v}}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_k} = \sum_{i=1}^{N}\left\{ \frac{d}{dt}\left(m_i\mathbf{v}_i \cdot \frac{\partial \mathbf{v}_i}{\partial \dot{q}_k}\right) - m_i\mathbf{v}_i \cdot \frac{\partial \mathbf{v}_i}{\partial q_k}\right\}$$

[11] Os deslocamentos virtuais podem ser entendidos como vetores tangentes ao espaço de configuração encarado como variedade diferenciável, evitando o uso de quantidades infinitesimais (Arnold 1976; José & Saletan 1998; Lemos 2005).

$$= \sum_{i=1}^{N} \left\{ \frac{d}{dt}\left[\frac{\partial}{\partial \dot{q}_k}\left(\frac{1}{2} m_i v_i^2\right)\right] - \frac{\partial}{\partial q_k}\left(\frac{1}{2} m_i v_i^2\right) \right\} = \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k}\right) - \frac{\partial T}{\partial q_k} \ , \qquad (1.4.11)$$

onde

$$T = \frac{1}{2} \sum_{i=1}^{N} m_i v_i^2 \qquad (1.4.12)$$

é a energia cinética do sistema. Podemos supor que a energia cinética T e as componentes Q_k da força generalizada estejam expressas exclusivamente em termos dos qs e $\dot{q}$s por meio das Eqs.(1.4.2) e (1.4.4). Levando (1.4.5), (1.4.7) e (1.4.11) em (1.3.9) somos conduzidos a

$$\sum_{k=1}^{n} \left\{ \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k}\right) - \frac{\partial T}{\partial q_k} - Q_k \right\} \delta q_k = 0 \ . \qquad (1.4.13)$$

Como os δqs são mutuamente independentes e arbitrários, esta última igualdade só pode ser satisfeita se o coeficiente de cada δq_k for zero. Inferimos, assim, as n equações

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_k}\right) - \frac{\partial T}{\partial q_k} = Q_k \quad , \quad k = 1, \ldots, n \ , \qquad (1.4.14)$$

às vezes chamadas de equações de Lagrange. Reservaremos esta designação, no entanto, para as situações mais importantes em que as forças aplicadas derivam de um potencial.

Na ausência de vínculos as Eqs.(1.4.14) permanecem válidas e representam a segunda lei de Newton expressa num sistema de coordenadas curvilíneas arbitrário, conforme passamos a ilustrar.

■ **Exemplo 1.4.3.** O movimento de uma partícula no espaço é descrito por meio de coordenadas cartesianas . O vetor posição da partícula é $\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$, de modo que as coordenadas generalizadas são $q_1 = x, q_2 = y, q_3 = z$ e as componentes da força generalizada são

$$Q_1 \equiv Q_x = \mathbf{F} \cdot \frac{\partial \mathbf{r}}{\partial x} = \mathbf{F} \cdot \hat{\mathbf{i}} = F_x \ , \ Q_2 = F_y \ , \ Q_3 = F_z \ .$$

A energia cinética é

$$T = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \ ,$$

donde:

$$\frac{\partial T}{\partial \dot{x}} = m\dot{x} \ , \ \frac{\partial T}{\partial \dot{y}} = m\dot{y} \ , \ \frac{\partial T}{\partial \dot{z}} = m\dot{z} \ , \ \frac{\partial T}{\partial x} = \frac{\partial T}{\partial y} = \frac{\partial T}{\partial z} = 0 \ .$$

As Eqs.(1.4.14) tomam a forma

$$m\ddot{x} = F_x \quad , \quad m\ddot{y} = F_y \quad , \quad m\ddot{z} = F_z$$

e recuperamos, assim, as equações de movimento newtonianas em sua forma original. ■

■ **Exemplo 1.4.4.** Uma partícula move-se num plano e coordenadas polares são empregadas para a descrição do movimento. O vetor posição da partícula escreve-se $\mathbf{r} = r\cos\theta\,\hat{\mathbf{i}} + r\,\mathrm{sen}\,\theta\,\hat{\mathbf{j}}$, e as componentes da força generalizada são

$$Q_1 \equiv Q_r = \mathbf{F} \cdot \frac{\partial \mathbf{r}}{\partial r} = \mathbf{F} \cdot (\cos\theta\,\hat{\mathbf{i}} + \mathrm{sen}\,\theta\,\hat{\mathbf{j}}) = \mathbf{F} \cdot \hat{\mathbf{e}}_r = F_r \ ,$$

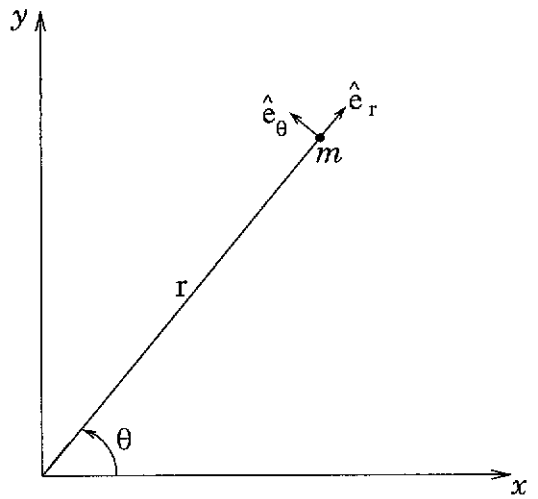


Figura 1.5: Partícula num plano em coordenadas polares.

$$Q_2 \equiv Q_\theta = \mathbf{F} \cdot \frac{\partial \mathbf{r}}{\partial \theta} = r\mathbf{F} \cdot (-\operatorname{sen}\theta\,\hat{\mathbf{i}} + \cos\theta\,\hat{\mathbf{j}}) = r\mathbf{F} \cdot \hat{\mathbf{e}}_\theta = rF_\theta \ ,$$

onde $\hat{\mathbf{e}}_r = \cos\theta\,\hat{\mathbf{i}} + \operatorname{sen}\theta\,\hat{\mathbf{j}}$ e $\hat{\mathbf{e}}_\theta = -\operatorname{sen}\theta\,\hat{\mathbf{i}} + \cos\theta\,\hat{\mathbf{j}}$ são os unitários radial e angular representados na Fig. 1.5. Usando

$$\dot{x} = \dot{r}\cos\theta - r\dot{\theta}\operatorname{sen}\theta \ , \quad \dot{y} = \dot{r}\operatorname{sen}\theta + r\dot{\theta}\cos\theta \ ,$$

a energia cinética expressa em termos de coordenadas polares é

$$T = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) \ .$$

Portanto, as Eqs.(1.4.14) assumem a forma

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{r}}\right) - \frac{\partial T}{\partial r} = Q_r \implies m\ddot{r} - mr\dot{\theta}^2 = F_r \ ,$$

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{\theta}}\right) - \frac{\partial T}{\partial \theta} = Q_\theta \implies \frac{d}{dt}(mr^2\dot{\theta}) = rF_\theta \ .$$

Verifica-se prontamente que rF_θ é a componente normal ao plano do movimento do torque em relação à origem, enquanto que $mr^2\dot{\theta}$ é a componente correspondente do momento angular. Desenvolvendo explicitamente a derivada temporal, as equações de movimento anteriores tornam-se

$$m\ddot{r} - mr\dot{\theta}^2 = F_r \quad , \quad mr\ddot{\theta} + 2m\dot{r}\dot{\theta} = F_\theta \ , \tag{1.4.15}$$

que são simplesmente as componentes polares da equação de movimento de Newton (Symon 1971). ■

Equações de Lagrange

As equações (1.4.14) adquirem uma forma particularmente concisa e elegante quando as forças $\mathbf{F}_i$ derivam de um potencial escalar $V(\mathrm{r}_1, \ldots, \mathrm{r}_N, t)$. Neste caso,

$$\mathbf{F}_i = -\boldsymbol{\nabla}_i V = -\Big(\frac{\partial V}{\partial x_i}\hat{\mathbf{i}} + \frac{\partial V}{\partial y_i}\hat{\mathbf{j}} + \frac{\partial V}{\partial z_i}\hat{\mathbf{k}}\Big) \tag{1.4.16}$$

e as forças generalizadas escrevem-se

$$Q_k = \sum_{i=1}^{N} \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_k} = -\sum_{i=1}^{N}\Big(\frac{\partial V}{\partial x_i}\frac{\partial x_i}{\partial q_k} + \frac{\partial V}{\partial y_i}\frac{\partial y_i}{\partial q_k} + \frac{\partial V}{\partial z_i}\frac{\partial z_i}{\partial q_k}\Big) = -\frac{\partial V}{\partial q_k} \ , \tag{1.4.17}$$

onde usamos a regra da cadeia da diferenciação. Com o emprego das Eqs.(1.4.2) o potencial V exprime-se como função exclusivamente dos qs, sendo independente das velocidades generalizadas. Inserindo (1.4.17) em (1.4.14) resulta

$$\frac{d}{dt}\Big(\frac{\partial T}{\partial \dot{q}_k}\Big) - \frac{\partial}{\partial q_k}(T - V) = 0 \ . \tag{1.4.18}$$

Dado que $\partial V/\partial \dot{q}_k = 0$, estas últimas equações são equivalentes a

$$\frac{d}{dt}\left[\frac{\partial}{\partial \dot{q}_k}(T - V)\right] - \frac{\partial}{\partial q_k}(T - V) = 0 \ . \tag{1.4.19}$$

Definindo a *função de Lagrange* ou, simplesmente, *lagrangiana* L por

$$L = T - V \ , \tag{1.4.20}$$

as equações de movimento do sistema podem ser escritas na forma

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_k}\Big) - \frac{\partial L}{\partial q_k} = 0 \quad , \quad k = 1, \ldots, n \ . \tag{1.4.21}$$

Doravante sempre nos referiremos às Eqs.(1.4.21) como *equações de Lagrange.*[12]

As Eqs.(1.4.21) constituem um sistema de n equações diferenciais ordinárias de segunda ordem que determinam univocamente os $q_k(t)$ desde que sejam dadas as $2n$ condições iniciais $q_1(t_0), \ldots, q_n(t_0)$ e $\dot{q}_1(t_0), \ldots, \dot{q}_n(t_0)$ num instante inicial t_0. Uma vez determinadas todas as coordenadas generalizadas em função do tempo, a posição de cada partícula do sistema em qualquer instante t determina-se por meio das Eqs.(1.4.2).

As equações de Lagrange fornecem o meio mais econômico de escrever as equações de movimento, pois envolvem o número mínimo de coordenadas, além de eliminar qualquer referência às forças de vínculo (o potencial V na lagrangiana (1.4.20) refere-se apenas às forças aplicadas). Em lugar das forças e acelerações vetoriais que caracterizam a abordagem newtoniana, no método de Lagrange basta lidar com duas funções escalares, T e V, o que introduz enormes simplificações no tratamento de problemas mecânicos. As equações de Lagrange possuem a vantagem adicional de serem válidas para uma escolha arbitrária das coordenadas generalizadas, a escolha em cada situação específica sendo ditada por razões de conveniência e simplicidade. É preciso salientar, no entanto, que a

[12] Nascido na Itália, em 1736, de uma família abastada de origem francesa pelo lado paterno, e falecido no ano de 1813, Joseph-Louis Lagrange, discípulo intelectual de Leonhard Euler, foi um dos maiores matemáticos de sua época. As equações de Lagrange – "Équations différentielles pour la solution de tous les problèmes de Dynamique" – vieram a lume em 1788 no seu grande livro *Mécanique Analytique*, no qual o tratamento adotado era puramente analítico, sem apelar para construções nem raciocínios geométricos. No prefácio Lagrange dizia, com orgulho, que sua obra não continha nenhum diagrama: "On ne trouvera point de Figures dans cet Ouvrage."

energia cinética T e a energia potencial V *devem ambas ser expressas relativamente a um mesmo referencial inercial.* Isso se deve ao fato de as equações de Lagrange (1.4.21) terem sido deduzidas a partir da segunda lei de Newton $\dot{\mathbf{p}}_i = \mathbf{F}_i$, que só vale em referenciais inerciais.

Invariância das Equações de Lagrange

Embora a invariância das equações de Lagrange sob uma transformação geral de coordenadas seja evidente da dedução feita acima, uma demonstração direta é instrutiva. Se $Q_1, \ldots, Q_n$ são novas cordenadas generalizadas que são funções arbitrárias das coordenadas generalizadas originais $q_1, \ldots, q_n$, temos[13]

$$Q_k = G_k(q_1, \ldots, q_n, t) \quad , \quad k = 1, \ldots, n \tag{1.4.22}$$

e, inversamente,

$$q_k = g_k(Q_1, \ldots, Q_n, t) \quad , \quad k = 1, \ldots, n \ . \tag{1.4.23}$$

A mudança de coordenadas (1.4.22) é conhecida como uma *transformação de ponto* porque mapeia pontos do espaço de configuração definido pelos qs em pontos do espaço de configuração definido pelos Qs. Na terminologia matemática, um mapeamento bijetivo diferenciável G com inverso $g = G^{-1}$ também diferenciável constitui um *difeomorfismo*, e o espaço de configuração dos Qs é dito *difeomorfo* ao espaço de configuração dos qs.

Diferenciando a Eq.(1.4.23) relativamente ao tempo, resulta

$$\dot{q}_k = \sum_{l=1}^{n} \frac{\partial q_k}{\partial Q_l} \dot{Q}_l + \frac{\partial q_k}{\partial t} \ , \tag{1.4.24}$$

donde

$$\frac{\partial \dot{q}_k}{\partial \dot{Q}_l} = \frac{\partial q_k}{\partial Q_l} \ . \tag{1.4.25}$$

A lagrangiana transformada $\bar{L}(Q, \dot{Q}, t)$ é simplesmente a lagrangiana original $L(q, \dot{q}, t)$ expressa em termos de $(Q, \dot{Q}, t)$:

$$\bar{L}(Q, \dot{Q}, t) = L(q(Q,t), \dot{q}(Q, \dot{Q}, t), t) \ . \tag{1.4.26}$$

Como $\partial q_k / \partial \dot{Q}_l = 0$, podemos escrever

$$\frac{\partial \bar{L}}{\partial \dot{Q}_i} = \sum_{k=1}^{n} \left(\frac{\partial L}{\partial q_k} \frac{\partial q_k}{\partial \dot{Q}_i} + \frac{\partial L}{\partial \dot{q}_k} \frac{\partial \dot{q}_k}{\partial \dot{Q}_i} \right) = \sum_{k=1}^{n} \frac{\partial L}{\partial \dot{q}_k} \frac{\partial q_k}{\partial Q_i} \ , \tag{1.4.27}$$

donde

$$\frac{d}{dt}\left(\frac{\partial \bar{L}}{\partial \dot{Q}_i} \right) = \sum_{k=1}^{n} \left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k} \right) \frac{\partial q_k}{\partial Q_i} + \frac{\partial L}{\partial \dot{q}_k} \frac{d}{dt}\left(\frac{\partial q_k}{\partial Q_i} \right) \right] = \sum_{k=1}^{n} \left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k} \right) \frac{\partial q_k}{\partial Q_i} + \frac{\partial L}{\partial \dot{q}_k} \frac{\partial \dot{q}_k}{\partial Q_i} \right] \ , \tag{1.4.28}$$

[13] Apesar da notação coincidente, estes Qs não devem ser confundidos com as componentes da força generalizada.

onde usamos a Eq.(1.4.9) com q_k no lugar de $\mathbf{r}_i$ e Q_i no lugar de q_k. Combinando esta última equação com

$$\frac{\partial \bar{L}}{\partial Q_i} = \sum_{k=1}^{n} \left(\frac{\partial L}{\partial q_k} \frac{\partial q_k}{\partial Q_i} + \frac{\partial L}{\partial \dot{q}_k} \frac{\partial \dot{q}_k}{\partial Q_i} \right) \tag{1.4.29}$$

resulta

$$\frac{d}{dt}\left(\frac{\partial \bar{L}}{\partial \dot{Q}_i} \right) - \frac{\partial \bar{L}}{\partial Q_i} = \sum_{k=1}^{n} \left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} \right] \frac{\partial q_k}{\partial Q_i} = 0 \ , \tag{1.4.30}$$

em virtude de (1.4.21), completando a demonstração. Em linguagem matemática: as equações de Lagrange são invariantes sob difeomorfismos.

Mesmo correndo o risco de aborrecer o leitor, repetimos que, embora possa ser expressa em termos de coordenadas generalizadas arbitrárias, a lagrangiana $L = T - V$ tem que ser escrita *inicialmente* em termos de posições e velocidades relativas a um referencial inercial.

1.5 Aplicações das Equações de Lagrange

O procedimento que deve ser seguido para se escrever as equações de Lagrange associadas a um dado sistema mecânico holônomo é simples. Em primeiro lugar precisam ser escolhidas coordenadas generalizadas $q_1, \ldots, q_n$. Em seguida, as energias cinética e potencial devem ser expressas em termos dos qs e $\dot{q}$s apenas, de tal modo que a lagrangiana $L = T - V$ também fique expressa somente em função das coordenadas e velocidades generalizadas. Basta, finalmente, calcular as derivadas parciais pertinentes de L, introduzi-las nas Eqs.(1.4.21), e está concluído o processo de construção das equações de movimento do sistema. Vejamos algumas ilustrações desse procedimento.

■ **Exemplo 1.5.1.** Obter uma lagrangiana e as respectivas equações de Lagrange para o sistema mecânico representado na Fig. 1.6, considerando desprezíveis as massas da roldana e do fio inextensível, e que o comprimento natural da mola é l.

Solução. Utilizando as coordenadas indicadas na Fig. 1.6 e supondo que o fio permanece sempre esticado, o vínculo $x_1 + x_2 = l_0$, onde l_0 é uma constante determinada pelo comprimento do fio e pelo raio da roldana, mostra que somente duas das três coordenadas x_1, x_2, x_3 podem ser tomadas como coordenadas generalizadas (o sistema só possui dois graus de liberdade). Escolhamos x_2 e x_3 como coordenadas generalizadas. A energia cinética do sistema é

$$T = \frac{m_1}{2}\dot{x}_1^2 + \frac{m_2}{2}\dot{x}_2^2 + \frac{m_3}{2}\dot{x}_3^2 = \frac{m_1 + m_2}{2}\dot{x}_2^2 + \frac{m_3}{2}\dot{x}_3^2 \ ,$$

porque de $x_1 = l_0 - x_2$ deduz-se $\dot{x}_1 = -\dot{x}_2$. Adotando o nível zero do potencial gravitacional no plano horizontal que passa pelo centro da polia, temos

$$V = -m_1 g x_1 - m_2 g x_2 - m_3 g x_3 + \frac{k}{2}(x_3 - x_2 - l)^2 = -(m_2 - m_1) g x_2 - m_1 g l_0 - m_3 g x_3 + \frac{k}{2}(x_3 - x_2 - l)^2 \ ,$$

pois pontos situados abaixo do nível zero têm potencial gravitacional negativo. A lagrangiana é

$$L = T - V = \frac{m_1 + m_2}{2}\dot{x}_2^2 + \frac{m_3}{2}\dot{x}_3^2 + (m_2 - m_1) g x_2 + m_3 g x_3 - \frac{k}{2}(x_3 - x_2 - l)^2 + m_1 g l_0 \ .$$

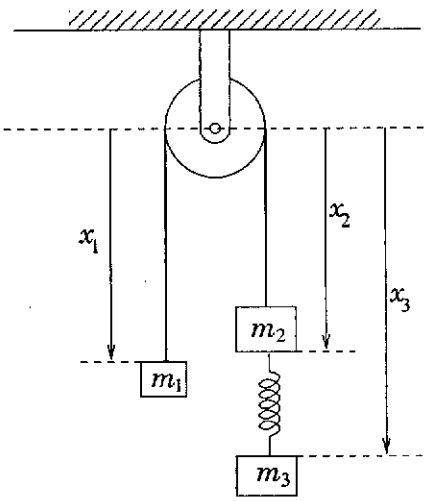


Figura 1.6: Sistema mecânico do Exemplo 1.5.1.

Temos

$$\frac{\partial L}{\partial \dot{x}_2} = (m_1 + m_2)\dot{x}_2 \quad , \quad \frac{\partial L}{\partial x_2} = (m_2 - m_1)g + k(x_3 - x_2 - l) \ ,$$
$$\frac{\partial L}{\partial \dot{x}_3} = m_3\dot{x}_3 \quad , \quad \frac{\partial L}{\partial x_3} = m_3 g - k(x_3 - x_2 - l) \ ,$$

e as equações de Lagrange são

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_2}\right) - \frac{\partial L}{\partial x_2} = 0 \quad \Longrightarrow \quad (m_1 + m_2)\ddot{x}_2 - (m_2 - m_1)g - k(x_3 - x_2 - l) = 0 \ ,$$
$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_3}\right) - \frac{\partial L}{\partial x_3} = 0 \quad \Longrightarrow \quad m_3\ddot{x}_3 - m_3 g + k(x_3 - x_2 - l) = 0 \ .$$

Se $k = 0$ não há interação entre m_2 e m_3. Neste caso limite as equações de Lagrange prevêem corretamente que m_3 cai em queda livre ($\ddot{x}_3 = g$) e que a aceleração $\ddot{x}_2 = (m_2 - m_1)g/(m_1 + m_2)$ da massa m_2 coincide com a obtida no tratamento da máquina de Atwood pelo princípio de d'Alembert (Exemplo 1.3.4). ■

■ **Exemplo 1.5.2.** Uma conta desliza ao longo de uma haste retilínea lisa que gira com velocidade angular constante num plano horizontal. Descrever o seu movimento pelo formalismo de Lagrange.

Solução. Seja xy o plano horizontal que contém a haste e usemos coordenadas polares para localizar a conta de massa m (ver Fig. 1.7). As variáveis r, θ não podem ser tomadas como coordenadas generalizadas porque θ é forçado a obedecer à restrição $\theta - \omega t = 0$, que é um vínculo holônomo da forma (1.2.7), onde ω é a velocidade angular constante da haste, suposta conhecida. O sistema possui somente um grau de liberdade (movimento radial) e podemos escolher $q_1 = r$ como coordenada generalizada. De acordo com o Exemplo 1.4.4, a energia cinética pode ser posta na forma

$$T = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) = \frac{m}{2}(\dot{r}^2 + \omega^2 r^2) \ ,$$

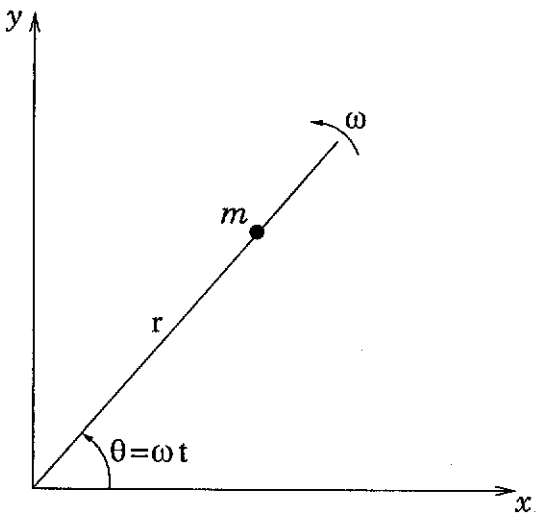


Figura 1.7: Conta deslizando ao longo de uma haste horizontal girante.

onde usamos $\dot{\theta} = \omega$. Adotando o nível zero do potencial gravitacional no plano do movimento, a lagrangiana do sistema reduz-se à energia cinética:

$$L = T - V = \frac{m}{2}(\dot{r}^2 + \omega^2 r^2) \ .$$

Dispondo da lagrangiana expressa exclusivamente em função de r e $\dot{r}$, a equação de movimento do sistema é imediatamente obtida:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = 0 \implies \frac{d}{dt}(m\dot{r}) - m\omega^2 r = 0 \implies \ddot{r} = \omega^2 r \ .$$

Conclui-se que a conta tende a afastar-se do eixo de rotação em conseqüência da "força centrífuga", que é o resultado bem conhecido. ■

■ **Exemplo 1.5.3.** Aplicar o formalismo lagrangiano para obter as equações de movimento de um pêndulo duplo plano.

Solução. Sejam (x_1, y_1) e (x_2, y_2) as coordenadas cartesianas das massas m_1 e m_2, respectivamente. Tomando os ângulos θ_1 e θ_2 como coordenadas generalizadas (vide Fig. 1.1 na página 12), temos

$$x_1 = l_1 \operatorname{sen}\theta_1 \ , \ y_1 = l_1 \cos\theta_1 \ , \ x_2 = l_1 \operatorname{sen}\theta_1 + l_2 \operatorname{sen}\theta_2 \ , \ y_2 = l_1 \cos\theta_1 + l_2 \cos\theta_2 \ ,$$

donde

$$\begin{aligned} \dot{x}_1 &= l_1\dot{\theta}_1 \cos\theta_1 \quad , \quad \dot{x}_2 = l_1\dot{\theta}_1 \cos\theta_1 + l_2\dot{\theta}_2 \cos\theta_2 \ , \\ \dot{y}_1 &= -l_1\dot{\theta}_1 \operatorname{sen}\theta_1 \quad , \quad \dot{y}_2 = -l_1\dot{\theta}_1 \operatorname{sen}\theta_1 - l_2\dot{\theta}_2 \operatorname{sen}\theta_2 \ . \end{aligned}$$

A energia cinética relativa ao referencial supostamente inercial (x, y) é

$$T = \frac{m_1}{2}(\dot{x}_1^2 + \dot{y}_1^2) + \frac{m_2}{2}(\dot{x}_2^2 + \dot{y}_2^2)$$

que, em termos das coordenadas e velocidades generalizadas, escreve-se

$$T = \frac{m_1 + m_2}{2} l_1^2 \dot{\theta}_1^2 + \frac{m_2}{2} l_2^2 \dot{\theta}_2^2 + m_2 l_1 l_2 \dot{\theta}_1 \dot{\theta}_2 \cos(\theta_1 - \theta_2) \ .$$

Por outro lado, com o nível zero do potencial gravitacional no plano horizontal que contém o ponto de suspensão de m_1, temos

$$V = -m_1 g y_1 - m_2 g y_2 = -(m_1 + m_2) g l_1 \cos\theta_1 - m_2 g l_2 \cos\theta_2 \ .$$

Finalmente, a lagrangiana $L = T - V$ é dada por

$$L = \frac{m_1 + m_2}{2} l_1^2 \, \dot{\theta}_1^2 + \frac{m_2 l_2^2}{2} \dot{\theta}_2^2 + m_2 l_1 l_2 \, \dot{\theta}_1 \, \dot{\theta}_2 \cos(\theta_1 - \theta_2) + (m_1 + m_2) g l_1 \, \cos\theta_1 + m_2 g l_2 \cos\theta_2 \, .$$

Usando

$$\begin{aligned} \frac{\partial L}{\partial \dot{\theta}_1} &= (m_1 + m_2) l_1^2 \, \dot{\theta}_1 + m_2 l_1 l_2 \, \dot{\theta}_2 \cos(\theta_1 - \theta_2) \ , \\ \frac{\partial L}{\partial \theta_1} &= -m_2 l_1 l_2 \, \dot{\theta}_1 \, \dot{\theta}_2 \, \mathrm{sen}\,(\theta_1 - \theta_2) - (m_1 + m_2) g l_1 \, \mathrm{sen}\,\theta_1 \ , \end{aligned}$$

a equação de Lagrange

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}_1}\right) - \frac{\partial L}{\partial \theta_1} = 0$$

toma a forma

$$(m_1 + m_2) l_1^2 \ddot{\theta}_1 + m_2 l_1 l_2 \ddot{\theta}_2 \cos(\theta_1 - \theta_2) + m_2 l_1 l_2 \dot{\theta}_2^2 \, \mathrm{sen}\,(\theta_1 - \theta_2) + (m_1 + m_2) g l_1 \, \mathrm{sen}\,\theta_1 = 0 \ .$$

De modo inteiramente análogo, obtém-se

$$m_2 l_2^2 \ddot{\theta}_2 + m_2 l_1 l_2 \ddot{\theta}_1 \cos(\theta_1 - \theta_2) - m_2 l_1 l_2 \dot{\theta}_1^2 \, \mathrm{sen}\,(\theta_1 - \theta_2) + m_2 g l_2 \, \mathrm{sen}\,\theta_2 = 0$$

como a segunda das equações de Lagrange. ■

1.6 Potenciais Generalizados e Função de Dissipação

As equações de Lagrange podem ter sua gama de validade estendida a certas situações em que as forças aplicadas dependem das velocidades generalizadas.

Potenciais Generalizados

Quando as forças generalizadas resultam de uma função $U(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t)$ por meio das expressões

$$Q_k = -\frac{\partial U}{\partial q_k} + \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{q}_k}\right) \ , \tag{1.6.1}$$

então as Eqs.(1.4.14) ainda implicam as Eqs.(1.4.21) com a lagrangiana definida por

$$L = T - U \ . \tag{1.6.2}$$

A função U é chamada de *potencial generalizado* ou *potencial dependente das velocidades.* A classe de forças abrangida pela Eq.(1.6.1) é mais ampla do que o conjunto das forças conservativas, estas últimas constituindo o caso particular em que U independe das velocidades generalizadas e do tempo. A admissão de forças dedutíveis de potenciais generalizados não é fruto de um desejo meramente acadêmico de buscar a máxima generalidade sem conseqüências físicas: a força eletromagnética sobre uma carga em movimento admite um potencial generalizado. Em virtude de sua enorme importância, este assunto merece uma discussão detalhada.

Lagrangiana de Partícula em Campo Eletromagnético Externo

A força experimentada por uma carga elétrica e em movimento num campo eletromagnético externo é a força de Lorentz (em unidades CGS gaussianas)

$$\mathbf{F} = e\Big(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B}\Big) \ . \tag{1.6.3}$$

As equações de Maxwell permitem escrever os campos em termos de um potencial escalar $\phi(\mathbf{r}, t)$ e de um potencial vetor $\mathbf{A}(\mathbf{r}, t)$ da seguinte maneira (Marion & Heald 1980):

$$\mathbf{E} = -\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \quad , \quad \mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A} \ . \tag{1.6.4}$$

Utilizando como coordenadas generalizadas as próprias coordenadas cartesianas da partícula, as componentes da força generalizada coincidem com as componentes cartesianas da força de Lorentz. Considere, portanto,

$$\mathbf{F} = e\left\{-\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} + \frac{1}{c}\mathbf{v} \times (\boldsymbol{\nabla} \times \mathbf{A})\right\} \ . \tag{1.6.5}$$

Pretendemos mostrar que $\mathbf{F}$ pode ser representada na forma (1.6.1) para alguma função U. Mas em (1.6.1) aparece uma derivada *total* em relação ao tempo, ao passo que em (1.6.5) a derivada é *parcial.* Podemos introduzir uma derivada total em (1.6.5) notando que

$$\frac{d\mathbf{A}}{dt} = \frac{\partial \mathbf{A}}{\partial x}\dot{x} + \frac{\partial \mathbf{A}}{\partial y}\dot{y} + \frac{\partial \mathbf{A}}{\partial z}\dot{z} + \frac{\partial \mathbf{A}}{\partial t} = (\mathbf{v}.\boldsymbol{\nabla})\,\mathbf{A} + \frac{\partial \mathbf{A}}{\partial t} \ . \tag{1.6.6}$$

Usando ainda

$$\mathbf{v} \times (\boldsymbol{\nabla} \times \mathbf{A}) = \boldsymbol{\nabla}(\mathbf{v}.\mathbf{A}) - (\mathbf{v}.\boldsymbol{\nabla})\,\mathbf{A} \ , \tag{1.6.7}$$

pois o operador nabla só afeta as variáveis de posição,[14] a Eq.(1.6.5) reduz-se a

$$\mathbf{F} = e\left\{-\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{d\mathbf{A}}{dt} + \frac{1}{c}\boldsymbol{\nabla}(\mathbf{v}.\mathbf{A})\right\} \ . \tag{1.6.8}$$

Com o uso do operador $\boldsymbol{\nabla}_v = \hat{\mathbf{i}}\partial/\partial\dot{x} + \hat{\mathbf{j}}\partial/\partial\dot{y} + \hat{\mathbf{k}}\partial/\partial\dot{z}$ e levando em conta que as coordenadas e velocidades generalizadas são tratadas como quantidades independentes, ficamos com

[14] Em geral vale a identidade $\boldsymbol{\nabla}(\mathbf{a}\cdot\mathbf{b}) = (\mathbf{a}\cdot\boldsymbol{\nabla})\mathbf{b} + (\mathbf{b}\cdot\boldsymbol{\nabla})\mathbf{a} + \mathbf{a}\times(\boldsymbol{\nabla}\times\mathbf{b}) + \mathbf{b}\times(\boldsymbol{\nabla}\times\mathbf{a})$.

$$\mathbf{F} = e\Big\{-\boldsymbol{\nabla}\big(\phi - \frac{1}{c}\mathbf{v}.\mathbf{A}\big)\Big\} - \frac{e}{c}\frac{d\mathbf{A}}{dt} = -\boldsymbol{\nabla}\big(e\phi - \frac{e}{c}\mathbf{v}.\mathbf{A}\big) + \frac{d}{dt}\Big[\boldsymbol{\nabla}_v\big(e\phi - \frac{e}{c}\mathbf{v}.\mathbf{A}\big)\Big] \ , \quad (1.6.9)$$

pois ϕ e $\mathbf{A}$ não dependem da velocidade.

Levando em conta que $q_1 = x, q_2 = y, q_3 = z$, a força $\mathbf{F}$ é da forma (1.6.1) com

$$U = e\phi - \frac{e}{c}\mathbf{v}.\mathbf{A} \ , \quad (1.6.10)$$

de modo que

$$L = T - U = \frac{mv^2}{2} - e\phi + \frac{e}{c}\mathbf{v}.\mathbf{A} \quad (1.6.11)$$

é a lagrangiana de uma partícula carregada num campo eletromagnético externo.

Função de Dissipação de Rayleigh

Se as forças generalizadas são da forma

$$Q_k = -\frac{\partial U}{\partial q_k} + \frac{d}{dt}\Big(\frac{\partial U}{\partial \dot{q}_k}\Big) + Q'_k \ , \quad (1.6.12)$$

onde Q'_k denota a parte das forças generalizadas que não provém de nenhum potencial generalizado, as equações de movimento (1.4.14) tornam-se

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_k}\Big) - \frac{\partial L}{\partial q_k} = Q'_k \ , \quad (1.6.13)$$

com $L = T - U$. Um caso de certa importância é aquele em que os Q'_k representam forças de atrito viscoso proporcionais às velocidades das partículas. Em componentes cartesianas,

$$F'_{ix} = -k_{ix}v_{ix} \ , \ F'_{iy} = -k_{iy}v_{iy} \ , \ F'_{iz} = -k_{iz}v_{iz} \ , \quad (1.6.14)$$

onde $\mathbf{F}'_i$ é a força dissipativa sobre a i-ésima partícula e k_{ix}, k_{iy}, k_{iz} são constantes positivas. Para facilitar um tratamento mais geral de tais situações, Rayleigh introduziu a chamada função de dissipação definida por

$$\mathcal{F} = \frac{1}{2}\sum_{i=1}^{N}(k_{ix}v_{ix}^2 + k_{iy}v_{iy}^2 + k_{iz}v_{iz}^2) \ , \quad (1.6.15)$$

de tal modo que

$$F'_{ix} = -\frac{\partial \mathcal{F}}{\partial v_{ix}} \ , \ F'_{iy} = -\frac{\partial \mathcal{F}}{\partial v_{iy}} \ , \ F'_{iz} = -\frac{\partial \mathcal{F}}{\partial v_{iz}} \ . \quad (1.6.16)$$

Considerando-se o trabalho realizado pelas forças dissipativas emerge o significado físico da função de dissipação de Rayleigh:

$$dW' = \sum_{i=1}^{N}\mathbf{F}'_i \cdot d\mathbf{r}_i = \sum_{i=1}^{N}\mathbf{F}'_i \cdot \mathbf{v}_i dt \implies \frac{dW'}{dt} = -\sum_{i=1}^{N}(k_{ix}v_{ix}^2 + k_{iy}v_{iy}^2 + k_{iz}v_{iz}^2) = -2\mathcal{F} \ ,$$

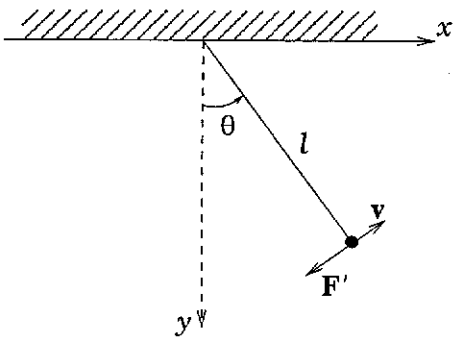


Figura 1.8: Pêndulo amortecido.

de sorte que $2\mathcal{F}$ é a taxa de dissipação da energia do sistema.

A parte dissipativa das forças generalizadas pode ser escrita como

$$Q'_k = \sum_{i=1}^{N} \mathbf{F}'_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_k} = \sum_{i=1}^{N} \mathbf{F}'_i \cdot \frac{\partial \mathbf{v}_i}{\partial \dot{q}_k} = -\sum_{i=1}^{N} \left(\frac{\partial \mathcal{F}}{\partial v_{ix}} \frac{\partial v_{ix}}{\partial \dot{q}_k} + \frac{\partial \mathcal{F}}{\partial v_{iy}} \frac{\partial v_{iy}}{\partial \dot{q}_k} + \frac{\partial \mathcal{F}}{\partial v_{iz}} \frac{\partial v_{iz}}{\partial \dot{q}_k} \right) = -\frac{\partial \mathcal{F}}{\partial \dot{q}_k} \ ,$$

onde usamos (1.4.10) e a regra da cadeia da diferenciação. Com este último resultado, as equações de movimento (1.6.13) transformam-se em

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} + \frac{\partial \mathcal{F}}{\partial \dot{q}_k} = 0 \ . \tag{1.6.17}$$

A principal vantagem – e, provavelmente, única – de se lançar mão de $\mathcal{F}$ é que ela permite escrever as equações de movimento na mesma forma (1.6.17) em qualquer sistema de coordenadas generalizadas.

■ **Exemplo 1.6.1.** Usar a função de dissipação de Rayleigh para descrever as oscilações de um pêndulo simples com resistência do ar proporcional à velocidade.

Solução. Usando coordenadas polares no plano xy da Fig. 1.8, temos $r = l$ e, de acordo com o Exemplo 1.4.4, $T = ml^2\dot{\theta}^2/2$. Como $V = -mgl\cos\theta$, resulta

$$L = \frac{ml^2}{2}\dot{\theta}^2 + mgl\cos\theta \ .$$

Por outro lado,

$$\mathcal{F} = \frac{1}{2}kv^2 = \frac{1}{2}kl^2\dot{\theta}^2 \ ,$$

supondo $k_x = k_y = k$. A equação (1.6.17) para θ toma a forma

$$\ddot{\theta} + \frac{k}{m}\dot{\theta} + \frac{g}{l}\operatorname{sen}\theta = 0 \ .$$

No caso de pequenas oscilações esta equação de movimento se reduz à de um oscilador harmônico amortecido. ■

1.7 Forças Centrais e Teorema de Bertrand

De todas as aplicações das leis do movimento, a mais célebre e notável é ao problema do movimento de uma partícula num potencial central, resolvido por Isaac Newton nos *Principia* no caso do potencial gravitacional.

Consideremos uma partícula em movimento sujeita a uma força $\mathbf{F}$ que depende somente da distância a um ponto fixo (o centro de força) e é dirigida ao longo da linha que une a partícula ao referido ponto, que tomaremos como origem de nosso sistema de coordenadas. Tendo em vista a simetria do problema, é conveniente empregar coordenadas esféricas r, θ, φ definidas por $x = r \operatorname{sen}\theta \cos\varphi, y = r \operatorname{sen}\theta \operatorname{sen}\varphi, z = r\cos\theta$. Em termos das coordenadas esféricas as componentes cartesianas da velocidade escrevem-se

$$\begin{aligned} \dot{x} &= \dot{r} \operatorname{sen}\theta \cos\varphi + r\dot{\theta}\cos\theta\cos\varphi - r\dot{\varphi}\operatorname{sen}\theta\operatorname{sen}\varphi \ , \\ \dot{y} &= \dot{r} \operatorname{sen}\theta \operatorname{sen}\varphi + r\dot{\theta}\cos\theta\operatorname{sen}\varphi + r\dot{\varphi}\operatorname{sen}\theta\cos\varphi \ , \\ \dot{z} &= \dot{r}\cos\theta - r\dot{\theta}\operatorname{sen}\theta \ , \end{aligned} \tag{1.7.1}$$

e a lagrangiana do sistema toma a forma

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - V(r) = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\,\dot{\varphi}^2\operatorname{sen}^2\theta) - V(r) \ , \tag{1.7.2}$$

onde $V(r)$ é o potencial da força central $\mathbf{F}$. A lagrangiana (1.7.2) pode ser obtida mais rapidamente lembrando que $\mathbf{v} = \dot{r}\,\hat{\mathbf{e}}_r + r\dot{\theta}\,\hat{\mathbf{e}}_\theta + r\,\dot{\varphi}\operatorname{sen}\theta\,\hat{\mathbf{e}}_\varphi$.

A força tem a mesma direção que o vetor posição $\mathbf{r}$, logo o torque em relação à origem é zero e o momento angular $\boldsymbol{l} = \mathbf{r}\times\mathbf{p}$ é um vetor constante. Como $\boldsymbol{l}$ é ortogonal ao plano definido por $\mathbf{r}$ e $\mathbf{p}$, o movimento se desenrola no plano perpendicular ao vetor $\boldsymbol{l}$. Escolhendo o eixo z paralelo a $\boldsymbol{l}$, o movimento ocorre no plano $z = 0$ ou $\theta = \pi/2$. Assim, r e φ tornam-se coordenadas polares no plano xy e, usando $\theta = \pi/2$ e $\dot{\theta} = 0$, a lagrangiana (1.7.2) toma a forma simplificada

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\varphi}^2) - V(r) \ . \tag{1.7.3}$$

As equações de Lagrange são

$$m\ddot{r} - mr\dot{\varphi}^2 + \frac{dV}{dr} = 0 \ , \tag{1.7.4}$$

$$\frac{d}{dt}(mr^2\dot{\varphi}) = 0 \ . \tag{1.7.5}$$

A segunda destas equações tem por solução

$$mr^2\dot{\varphi} = l = \text{constante} \ , \tag{1.7.6}$$

que corresponde à lei das áreas de Kepler: a linha que une o Sol a cada planeta varre áreas iguais em tempos iguais. A constante l é o módulo do momento angular.

Pela lei da conservação da energia,

$$T + V = \frac{m}{2}(\dot{r}^2 + r^2\dot{\varphi}^2) + V(r) = E = \text{constante} \ . \tag{1.7.7}$$

Com o uso de (1.7.6) esta equação pode se posta na forma

$$E = \frac{m}{2}\dot{r}^2 + \frac{l^2}{2mr^2} + V(r) \equiv \frac{m}{2}\dot{r}^2 + V_{\text{ef}}(r) \ , \tag{1.7.8}$$

onde

$$V_{\text{ef}}(r) = \frac{l^2}{2mr^2} + V(r) \ . \tag{1.7.9}$$

A Eq.(1.7.8) mostra que o movimento radial é equivalente ao movimento unidimensional de uma partícula no potencial efetivo $V_{\text{ef}}(r)$.

■ **Exercício 1.7.1.** Se $r = r_0$ é uma órbita circular com energia E_0, prove que $E_0 = V_{\text{ef}}(r_0)$ e $V'_{\text{ef}}(r_0) = 0$ ou, equivalentemente, $l^2/mr_0^3 = V'(r_0)$. Prove, ainda, que

$$V''_{\text{ef}}(r_0) = \frac{3}{r_0}V'(r_0) + V''(r_0) \ . \tag{1.7.10}$$

Sugestão: use as equações (1.7.4), (1.7.6) e (1.7.8). ■

Com ajuda de

$$\frac{d}{dt} = \frac{d\varphi}{dt}\frac{d}{d\varphi} = \dot{\varphi}\frac{d}{d\varphi} = \frac{l}{mr^2}\frac{d}{d\varphi} \tag{1.7.11}.$$

pode-se eliminar o tempo da Eq.(1.7.8), que toma a forma

$$\frac{l^2}{2mr^4}\left(\frac{dr}{d\varphi}\right)^2 + V_{\text{ef}}(r) = E \ , \tag{1.7.12}$$

e da qual deduz-se a equação da trajetória:

$$\varphi = \int \frac{l\,dr}{r^2\sqrt{2m[E - V_{\text{ef}}(r)]}} \ . \tag{1.7.13}$$

Problema de Kepler

No caso em que a força é de natureza gravitacional, $V(r) = -\kappa/r$ com $\kappa = GmM$, onde m é a massa do planeta e M é a massa do Sol. Com a introdução da variável $u = 1/r$ a integração em (1.7.13) torna-se elementar:

$$\varphi = -\int \frac{l\,du}{\sqrt{2mE + 2m\kappa u - l^2u^2}} = \cos^{-1}\frac{l/r - m\kappa/l}{\sqrt{2mE + \kappa^2m^2/l^2}} \ , \tag{1.7.14}$$

onde a origem do ângulo foi escolhida de modo a anular a constante de integração. Definindo

$$e = \sqrt{1 + \frac{2El^2}{m\kappa^2}} \ , \tag{1.7.15}$$

a solução de (1.7.14) para r é

$$\frac{1}{r} = \frac{m\kappa}{l^2}(1 + e\cos\varphi) \ . \tag{1.7.16}$$

Esta equação representa uma seção cônica tendo como foco a origem do sistema de coordenadas. Se $E < 0$ então $e < 1$ e a órbita é uma elipse de excentricidade e. Se $E = 0\,(e = 1)$ a órbita é uma parábola, e se $E > 0\,(e > 1)$ a órbita é uma hipérbole. Em Goldstein(1980) o leitor encontrará uma discussão mais ampla e detalhada do problema do movimento sob forças centrais em geral, e do problema de Kepler em particular.

■ **Exercício 1.7.2.** (i) Mostre que os valores máximo e mínimo de r são $r_{\max} = p(1-e)^{-1}$ e $r_{\min} = p(1+e)^{-1}$, onde $p = l^2/m\kappa$, e conclua daí que o semi-eixo maior da elipse $a = (r_{\max} + r_{\min})/2$ pode ser expresso como $a = p(1-e^2)^{-1}$. (ii) Se o ângulo φ cresce de $d\varphi$, a área varrida pelo raio vetor r é $dS = \frac{1}{2}r^2 d\varphi$. Mostre que $dS/dt = l/2m$. Num período τ do movimento, a área varrida pelo raio vetor é igual à área $S = \pi ab$ da elipse. Sabendo que, por definição de excentricidade, o semi-eixo menor da elipse é $b = a\sqrt{1-e^2}$, prove que $\tau = 2\pi a^{3/2}\sqrt{m/\kappa} = 2\pi a^{3/2}/\sqrt{GM}$. Desprezando o movimento do Sol, esta é a terceira lei de Kepler, segundo a qual o quadrado do período é proporcional ao cubo do semi-eixo maior com a mesma constante de proporcionalidade para todos os planetas. ■

Teorema de Bertrand

Uma órbita é dita limitada se a distância da partícula ao centro de força é constante (órbita circular) ou está compreendida entre dois valores extremos, isto é, $r_{\min} \leq r \leq r_{\max}$. Uma órbita limitada não precisa ser fechada: a partícula pode dar voltas indefinidamente sempre dentro do anel $r_{\min} \leq r \leq r_{\max}$ sem que a trajetória jamais se feche. Como acabamos de ver, no caso do potencial gravitacional ou de um potencial coulombiano atrativo todas as órbitas limitadas são fechadas. Isto é uma acontecimento excepcional para potenciais centrais, como mostra um resultado obtido por J. Bertrand em 1873.

Teorema de Bertrand. Os únicos potenciais centrais para os quais todas as órbitas limitadas são fechadas são $V(r) = -\kappa/r$ e $V(r) = \kappa r^2$, com $\kappa > 0$, isto é, ou a força é inversamente proporcional ao quadrado da distância ou obedece à lei de Hooke.

Demonstração. Seja φ o deslocamento angular ocorrido durante uma oscilação radial completa, de $r = r_{\min}$ até $r = r_{\max}$ e de volta a $r = r_{\min}$. Uma órbita limitada é fechada se, depois de um número inteiro de oscilações completas de r, o ângulo de rotação varia de um número inteiro de 2π radianos: $n\varphi = m2\pi$. Pela equação (1.7.13),

$$\varphi = 2\int_{r_{\min}}^{r_{\max}} \frac{l\,dr}{r^2\sqrt{2m[E - V_{\text{ef}}(r)]}} = 2\int_{u_{\min}}^{u_{\max}} \frac{l\,du}{\sqrt{2m[E - W(u)]}} \ , \tag{1.7.17}$$

onde $u = 1/r$ e

$$W(u) = V_{\text{ef}}\Big(\frac{1}{u}\Big) = \frac{l^2u^2}{2m} + V\Big(\frac{1}{u}\Big) \ . \tag{1.7.18}$$

Se todas as órbitas limitadas são fechadas, as que são vizinhas a trajetórias circulares são igualmente fechadas. Como indica o problema unidimensional equivalente definido pela Eq.(1.7.8), o potencial efetivo $V_{\text{ef}}(r)$ passa por um mínimo num raio r_0 compreendido entre $r_{\min}$ e $r_{\max}$. Consideremos uma órbita limitada resultante de uma ligeira perturbação da órbita circular de raio r_0 e energia E_0. Escrevendo $E = E_0 + \Delta E$ e expandindo $W(u)$ em torno de $u = u_0 = 1/r_0$ incluindo até termos de segunda ordem, resulta

$$E - W(u) = E_0 + \Delta E - W(u_0) - W'(u_0)(u - u_0) - \frac{1}{2}W''(u_0)(u - u_0)^2 \ . \tag{1.7.19}$$

De acordo com o Exercício 1.7.1, temos $E_0 - W(u_0) = E_0 - V_{\text{ef}}(r_0) = 0$ e, ainda,

$$W'(u_0) = \left[\frac{l^2u}{m} - \frac{1}{u^2}V'\Big(\frac{1}{u}\Big)\right]_{u=1/r_0} = \frac{1}{r_0}\left[\frac{l^2}{m} - r_0^3V'(r_0)\right] = 0 \ , \tag{1.7.20}$$

$$W''(u_0) = \left[\frac{l^2}{m} + \frac{2}{u^3}V'\Big(\frac{1}{u}\Big) + \frac{1}{u^4}V''\Big(\frac{1}{u}\Big)\right]_{u=1/r_0} = 3r_0^3V'(r_0) + r_0^4V''(r_0) = r_0^4\,V''_{\text{ef}}(r_0) \ . \tag{1.7.21}$$

Conseqüentemente,

$$E - W(u) = \Delta E - \frac{1}{2}W''(u_0)(u - u_0)^2 \tag{1.7.22}$$

e a expressão (1.7.17) para o ângulo de rotação se reduz a

$$\varphi = 2l\int_{u_{\min}}^{u_{\max}} \frac{du}{\sqrt{2m\Delta E - mW''(u_0)(u - u_0)^2}} \ . \tag{1.7.23}$$

Fazendo $a^2 = 2\Delta E/W''(u_0)$ e $u = u_0 + a\,\text{sen}\,\theta$ resulta

$$\varphi = \frac{2l}{\sqrt{mW''(u_0)}}\int_{-\pi/2}^{\pi/2} \frac{a\cos\theta d\theta}{\sqrt{a^2(1 - \text{sen}^2\,\theta)}} = \frac{2\pi l}{r_0^2\sqrt{mV''_{\text{ef}}(r_0)}} = 2\pi\sqrt{\frac{V'(r_0)}{3V'(r_0) + r_0V''(r_0)}} \ , \tag{1.7.24}$$

onde voltamos a fazer uso dos resultados do Exercício 1.7.1. Variando o momento angular l o raio r_0 varia continuamente, mas φ não pode variar continuamente porque é um múltiplo racional de 2π. O impasse é resolvido impondo que o termo mais à direita na cadeia de igualdades Eq.(1.7.24) seja uma constante independente de r_0, o que exige que o potencial $V(r)$ satisfaça a equação diferencial $rV''(r) + 3V'(r) = \alpha V'(r)$, onde α é uma constante positiva. Portanto, os potenciais admissíveis são apenas $V(r) = \kappa r^\beta$ ($\beta > -2$, $\beta \neq 0$) ou $V(r) = \kappa \ln r$. Para esses potenciais,

$$\varphi = \frac{2\pi}{\sqrt{\beta+2}} \ , \tag{1.7.25}$$

com $\beta = 0$ correspondendo ao caso logarítmico. O potencial logarítmico não serve porque $2\pi/\sqrt{2}$ não é múltiplo racional de 2π. O resultado (1.7.25) não depende da energia, logo deve permanecer válido nos limites $E \to \infty$ ou $E \to 0$. A idéia essencial do restante da demonstração consiste em obter expressões para o ângulo de rotação nesse casos especiais e exigir que elas coincidam com (1.7.25).

Como o potencial tem que ser atrativo para que existam órbitas limitadas, $V'(r) > 0$. Portanto, $\kappa > 0$ se $\beta > 0$ e $\kappa < 0$ se $-2 < \beta < 0$. No caso $V(r) = \kappa r^\beta$ com $\kappa > 0$ e $\beta > 0$, consideremos o limite $E \to \infty$. Da equação $E = l^2/2mr^2 + \kappa r^\beta$, válida para $r = r_{\min}$ ou $r = r_{\max}$, decorre que $r_{\max} \to \infty$ e $r_{\min} \to 0$ no limite $E \to \infty$, de modo que $u_{\min} \to 0$ e $u_{\max} \to \infty$. Com a mudança de variáveis $u = yu_{\max}$ a expressão (1.7.17) para o ângulo de rotação transforma-se em

$$\varphi = \frac{2l}{\sqrt{2m}} \int_{y_{\min}}^{1} \frac{dy}{\sqrt{\mathcal{U}(1) - \mathcal{U}(y)}} \ , \tag{1.7.26}$$

onde

$$\mathcal{U}(y) = y^2 \left[\frac{l^2}{2m} + \frac{\kappa}{(yu_{\max})^{\beta+2}} \right] \ . \tag{1.7.27}$$

No limite $E \to \infty$ a integral (1.7.26) reduz-se a

$$\varphi = \frac{2l}{\sqrt{2m}} \int_0^1 \frac{dy}{\sqrt{\frac{l^2}{2m} - \frac{l^2 y^2}{2m}}} = \frac{2l}{\sqrt{2m}} \frac{\sqrt{2m}}{l} \int_0^1 \frac{dy}{\sqrt{1-y^2}} = \pi \ . \tag{1.7.28}$$

Comparando este resultado com (1.7.25) resulta $\beta = 2$.

No caso de κ e β negativos, é melhor escrever $\kappa = -\sigma$ e $\beta = -\lambda$ com $\sigma > 0$ e $0 < \lambda < 2$. No limite $E \to 0$ os valores extremos de u são as soluções de $0 = \frac{l^2}{2mr^2} - \frac{\sigma}{r^\lambda} = \frac{l^2 u^2}{2m} - \sigma u^\lambda$, a saber, $u_{\min} = 0$ e $u_{\max} = (2m\sigma/l^2)^{1/(2-\lambda)}$. Com a mesma substituição anterior $u = yu_{\max}$ e $E = 0$ a expressão (1.7.17) para o ângulo de rotação torna-se

$$\varphi = \frac{2l}{\sqrt{2m}} \int_0^1 \frac{dy}{\sqrt{-\frac{l^2 y^2}{2m} + \frac{\sigma y^\lambda}{u_{\max}^{2-\lambda}}}} = 2 \int_0^1 \frac{dy}{\sqrt{y^\lambda - y^2}} \ . \tag{1.7.29}$$

A substituição $x = y^{(2-\lambda)/2}$ conduz a

$$\varphi = \frac{4}{2-\lambda} \int_0^1 \frac{dx}{\sqrt{1-x^2}} = \frac{2\pi}{2-\lambda} = \frac{2\pi}{2+\beta} \ . \tag{1.7.30}$$

Comparando este resultado com (1.7.25) resulta $\beta = -1$, o que completa a demonstração do teorema. □

A demonstração acima seguiu as indicações de Arnold(1976). Para provas alternativas, mas bem menos elegantes, o leitor é remetido a Goldstein(1980) ou Desloge(1982).

Note que $\varphi = 2\pi$ no caso do potencial inversamente proporcional ao quadrado da distância, de modo que numa única oscilação radial a órbita se fecha. No caso da força

que obedece à lei de Hooke, $\varphi = \pi$ e a órbita só se fecha depois de duas oscilações radiais completas.

A distâncias planetárias, a teoria geral da relatividade introduz uma diminuta correção proporcional a $1/r^3$ ao potencial newtoniano do Sol. Como previsto pelo teorema de Bertrand, as órbitas dos planetas não são fechadas: cada planeta descreve uma elipse cujo eixo maior gira lentamente em torno do Sol. O efeito é mais pronunciado para o planeta Mercúrio. A previsão correta do valor da taxa de precessão do periélio de Mercúrio é um dos maiores triunfos da teoria da gravitação de Einstein.

PROBLEMAS

1.1. A discussão do Exemplo 1.2.7 é incompleta por uma razão sutil: para decidir se um sistema é ou não holônomo precisamos considerar todos os vínculos ao mesmo tempo, pois os vínculos em conjunto podem ser integráveis embora cada um deles tomado isoladamente não seja integrável (Neĭmark & Fufaev 1972). Por exemplo, considere os vínculos

$$(x^2+y^2)dx + xzdz = 0 \quad , \quad (x^2+y^2)dy + yzdz = 0 \ .$$

Cada um deles separadamente é não-integrável. (i) Prove, no entanto, que juntos eles equivalem a

$$xd(x^2+y^2+z^2) = 0 \quad , \quad d\ln\frac{y}{x} = 0 \ .$$

(ii) Conclua que o sistema é holônomo com vínculos

$$x^2+y^2+z^2 = C_1 \quad , \quad y = C_2 x \ .$$

1.2. Uma conta de massa m desliza sem atrito ao longo de uma haste rígida, de massa desprezível, que gira num plano vertical com velocidade angular constante ω. Mostre que, com uma escolha adequada da coordenada r, a lagrangiana do sistema é

$$L = \frac{m}{2}\dot{r}^2 + \frac{m\omega^2}{2}r^2 - mgr\,\mathrm{sen}\,\omega t \ .$$

Dadas as condições iniciais $r(0) = r_0$, $\dot{r}(0) = 0$, encontre a solução da equação de Lagrange para r.

1.3. Obtenha a lagrangiana e as equações de Lagrange para um pêndulo esférico, isto é, uma massa m suspensa por um fio leve e inextensível de comprimento l cujo movimento não está restrito a um plano.

1.4. O pêndulo cicloidal de Huyghens consiste numa partícula oscilando num plano vertical ao longo de um arco de ciclóide com equações paramétricas

$$x = R\theta + R\,\mathrm{sen}\,\theta \quad , \quad y = -R\cos\theta \ .$$

Mostre que a lagrangiana desse sistema pode ser posta na forma

$$L = 2mR^2\cos^2\frac{\theta}{2}\,\dot{\theta}^2 + mgR\cos\theta \ .$$

Fazendo a transformação de ponto $u = \mathrm{sen}\,(\theta/2)$, obtenha a lagrangiana e a equação de Lagrange em termos da coordenada u. Prove que o período de oscilação é igual a $4\pi(R/g)^{1/2}$, independente, portanto, da amplitude θ_0.

1.5. Considere a chamada máquina de Atwood oscilante. Usando as coordenadas indicadas na figura, mostre que a lagrangiana é dada por

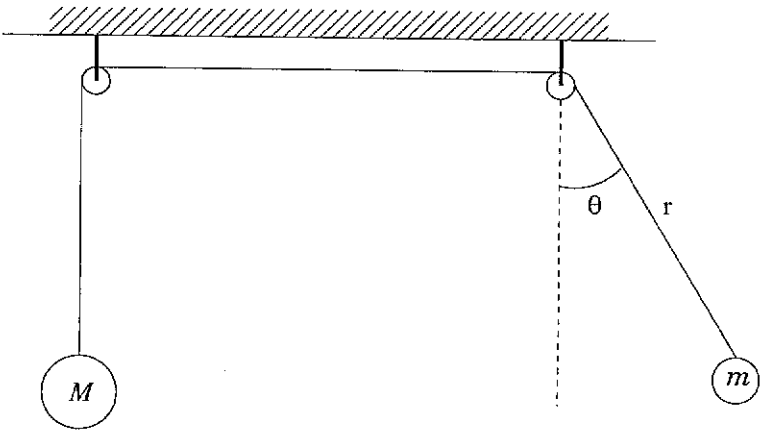


Figura 1.9: Máquina de Atwood oscilante do Problema 1.5 (M só se move verticalmente).

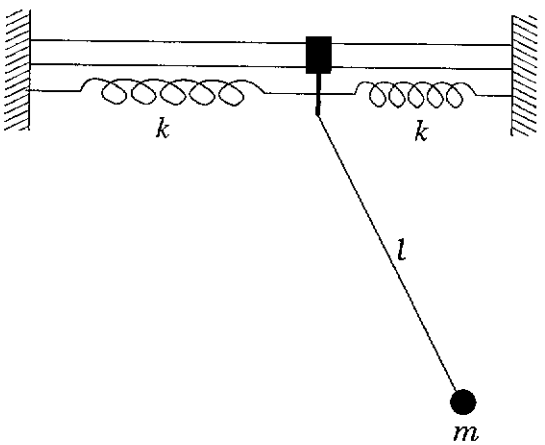


Figura 1.10: Problema 1.8.

$$L = \frac{m+M}{2}\dot{r}^2 + \frac{m}{2}r^2\dot{\theta}^2 - gr(M - m\cos\theta)$$

e escreva as equações de Lagrange.

1.6. Uma partícula cai verticalmente sob a ação da gravidade. Supondo que a força de resistência do ar seja proporcional à velocidade, obtenha a equação de movimento da partícula com a ajuda da função de dissipação $\mathcal{F} = \lambda v^2/2$. Determine a velocidade como função do tempo e prove que a maior velocidade possível para queda a partir do repouso é $v = mg/\lambda$.

1.7. Certos sistemas dissipativos simples admitem uma formulação lagrangiana que dispensa o emprego da função de dissipação de Rayleigh. Mostre que a equação de movimento do problema anterior pode ser obtida da lagrangiana

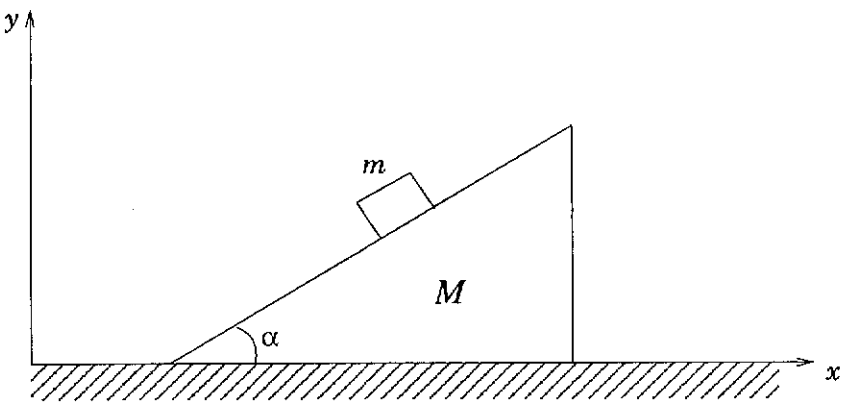


Figura 1.11: Problema 1.12.

$$L = e^{\lambda t}\left(\frac{m}{2}\dot{x}^2 - mgx\right) .$$

1.8. O ponto de suspensão de um pêndulo de massa m e comprimento l pode mover-se horizontalmente ligado a duas molas idênticas de constante elástica k. (i) Escolhendo como coordenadas generalizadas o deslocamento x do ponto de suspensão e o ângulo θ entre o fio e a vertical, encontre a lagrangiana e as equações de movimento. (ii) Obtenha a forma aproximada das equações de Lagrange no caso de pequenas oscilações angulares ($\operatorname{sen}\theta \approx \theta$, $\cos\theta \approx 1$ e termos do tipo θ^n ou $\dot{\theta}^n$ com $n > 1$ são desprezíveis). Combinando as equações de movimento aproximadas, mostre que $x = \alpha\theta$ para um certo valor da constante α. Mostre, em seguida, que o sistema equivale a um pêndulo simples de comprimento l' e exprima l' em termos de l, m, g, k.

1.9. Na eletrodinâmica de Weber a força entre duas cargas elétricas em movimento é dirigida ao longo da linha que as une e tem magnitude

$$F = \frac{ee'}{r^2}\left\{1 + \frac{r\ddot{r}}{c^2} - \frac{\dot{r}^2}{2c^2}\right\} ,$$

onde r denota a distância entre as partículas carregadas e c é a velocidade da luz no vácuo. Determine o potencial generalizado $U(r, \dot{r})$ associado a essa força. Escreva a lagrangiana e as equações de Lagrange para o movimento num plano de uma carga na presença de outra carga fixa na origem do sistema de coordenadas. [Embora conduza à expressão correta para a lei de Faraday e para a indutância de circuitos elétricos, a força de Weber implica que em certas situações uma carga se comporta como se sua massa fosse negativa (Maxwell 1891; Whittaker 1951).]

1.10. Um pêndulo elástico consiste numa massa m capaz de oscilar num plano vertical suspensa por uma mola de constante elástica k e comprimento natural l. Escolhendo coordenadas generalizadas convenientes, obtenha a lagrangiana e as equações de Lagrange.

1.11. Um pêndulo de massa m e comprimento l está suspenso por um ponto que executa oscilações horizontais da forma $x = a\cos\omega t$. Obtenha a lagrangiana em termos

do ângulo θ que o fio faz com a direção vertical orientada para baixo. Escreva a equação de movimento e mostre que, para pequenos valores de θ, ela reduz-se à de um oscilador harmônico forçado. Encontre uma solução estacionária da forma $\theta = \theta_0 \cos \omega t$. Como a amplitude θ_0 depende de m, l, a e ω? O que acontece quando $\omega = (g/l)^{1/2}$?

1.12. Um bloco de massa m desliza sem atrito ao longo de uma cunha de massa M e ângulo α. A cunha, por sua vez, escorrega sem atrito sobre uma superfície horizontal. (i) Escolhendo coordenadas generalizadas convenientes, obtenha a lagrangiana e as equações de movimento do sistema. (ii) Determine a aceleração da cunha relativamente ao sistema de referência inercial (x, y). (iii) Determine a aceleração do bloco em relação à cunha.

1.13. Uma partícula num campo gravitacional uniforme $\mathbf{g} = -g\hat{\mathbf{k}}$ está restrita a mover-se na superfície de um paraboloide de revolução definido pela equação $z = \rho^2/R$, onde R é uma constante positiva e (ρ, θ, z) são coordenadas cilíndricas. Obtenha a lagrangiana e as equações de Lagrange. Prove que a projeção do raio vetor da partícula sobre o plano xy varre áreas iguais em tempos iguais.

1.14. Prove que a lagrangiana de Bateman (1931)

$$L = e^{\lambda t}\left(\frac{m}{2}\dot{x}^2 - \frac{m\omega^2}{2}x^2\right)$$

gera a equação de movimento de um oscilador harmônico amortecido. Mostre que a transformação de ponto $q = e^{\lambda t/2}x$ elimina a dependência temporal explícita da lagrangiana.

1.15. Demonstre que a lagrangiana

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\varphi}^2 \operatorname{sen}^2\theta) - \frac{eg}{c}\dot{\varphi}\cos\theta$$

descreve o movimento de uma partícula carregada no campo $\mathbf{B} = g\mathbf{r}/r^3$ de um monopolo magnético e obtenha as equações de Lagrange. Sugestão: verifique que o potencial vetor de um monopolo magnético em coordenadas esféricas tem componentes $A_r = 0$, $A_\theta = 0$ e $A_\varphi = g(1 - \cos\theta)/r \operatorname{sen}\theta$.

1.16. No problema de Kepler, mostre que, para órbitas elípticas, a relação entre r e t pode ser posta na forma

$$t = \sqrt{\frac{m}{2}}\int_{r_{min}}^{r} \frac{dr}{\sqrt{E + \frac{\kappa}{r} - \frac{l^2}{2mr^2}}} = \sqrt{\frac{m}{2\kappa}}\int_{r_{min}}^{r} \frac{r\,dr}{\sqrt{r - \frac{r^2}{2a} - \frac{a(1-e^2)}{2}}}$$

onde a é o semi-eixo maior, e é a excentricidade e $t = 0$ foi escolhido no momento de passagem do planeta pelo periélio. Mostre que, em termos do ângulo ψ conhecido como anomalia excêntrica e definido por $r = a(1 - e\cos\psi)$, tem-se

$$\omega t = \psi - e \operatorname{sen}\psi \ ,$$

com $\omega = \sqrt{\kappa/ma^3}$. Esta última equação transcendente, que determina implicitamente ψ como função de t, é conhecida como *equação de Kepler*.

Capítulo 2

PRINCÍPIO VARIACIONAL DE HAMILTON

> Desde que existe como ciência, a Física tem como seu objetivo mais cobiçado a solução do problema de condensar todos os fenômenos naturais num único princípio. Dentre as leis mais ou menos gerais que marcam as conquistas da ciência física durante o curso dos últimos séculos, o princípio da mínima ação é talvez aquele que, no que se refere à forma e ao conteúdo, mais se aproxima desse objetivo final da pesquisa teórica.
>
> Max Planck

De todas a formulações da dinâmica clássica, a mais concisa tem a forma de um princípio variacional. Conforme idéias típicas do século XVIII, das quais Maupertuis foi um dos pioneiros, dentre todas as alternativas à sua disposição a natureza segue o curso mais econômico de acordo com algum critério de comparação entre as diversas possibilidades. O princípio diferencial de d'Alembert, do qual se deduzem as equações de Lagrange, exprime a lei fundamental do movimento em termos da configuração instantânea do sistema e de desvios infinitesimais da referida configuração. É possível reformular a lei dinâmica fundamental como um princípio integral, que leva em conta o movimento completo do sistema durante um intervalo de tempo finito. O princípio de Hamilton reduz as leis da mecânica a um enunciado segundo o qual, comparado com todos os movimentos imagináveis, o movimento real é aquele para o qual é mínima (mais geralmente, estacionária) uma certa quantidade – a *ação* – cujo valor depende do movimento do sistema em sua totalidade. A formulação precisa do *princípio de Hamilton* ou *princípio da mínima ação* requer uma breve incursão por um importante ramo da Matemática conhecido como cálculo das variações.

2.1 Rudimentos do Cálculo das Variações

O cálculo das variações ocupa-se com o problema de determinar extremos, isto é, máximos ou mínimos, de funcionais. Um funcional é uma função real cujo domínio é um

espaço de funções. Mais explicitamente, um funcional associa um número real a cada função de um certa classe de funções para as quais o funcional está definido.

■ **Exemplo 2.1.1.** Sejam (x_1, y_1) e (x_2, y_2) dois pontos distintos num plano, com $x_2 > x_1$. Seja $y(x)$ uma curva ligando esses dois pontos. Uma vez que o comprimento de arco infinitesimal no plano é

$$ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \ ,$$

o comprimento de arco da curva entre os pontos inicial e final é

$$s[y] = \int_{x_1}^{x_2} \sqrt{1 + y'(x)^2}\, dx \ . \tag{2.1.1}$$

O comprimento de arco s é um funcional de y, isto é, a cada função continuamente diferenciável $y(x)$ corresponde o único número real $s[y]$ definido pela Eq.(2.1.1). ■

O mais simples de todos os problemas do cálculo das variações pode ser assim formulado: sendo dado o funcional

$$J[y] = \int_{x_1}^{x_2} f(y(x), y'(x), x)\, dx \ , \tag{2.1.2}$$

onde $f : \boldsymbol{R}^3 \to \boldsymbol{R}$ é uma função conhecida, encontrar, entre todas as curvas continuamente diferenciáveis $y(x)$ que passam pelos pontos fixos (x_1, y_1) e (x_2, y_2), aquela que minimiza (mais geralmente, extremiza) J.

A fim de encontrar a curva que extremiza J, lançaremos mão de um artifício que reduz o problema ao de achar pontos extremos de uma função real de uma variável real. Seja $y(x)$ a função procurada[1] que extremiza J e consideremos uma curva vizinha $\overline{y}$ definida por (Fig. 2.1)

$$\overline{y}(x) = y(x) + \epsilon\eta(x) \ , \tag{2.1.3}$$

onde ϵ é um parâmetro real arbitrário e $\eta(x)$ é uma função continuamente diferenciável que se anula em $x = x_1$ e $x = x_2$:

$$\eta(x_1) = \eta(x_2) = 0 \ . \tag{2.1.4}$$

Estas últimas condições são necessárias para que a curva variada $\overline{y}$ também passe pelos pontos extremos (x_1, y_1) e (x_2, y_2). Substituindo y por $\overline{y}$ em (2.1.2) o resultado é uma função de ϵ:

$$\Phi(\epsilon) \equiv J[\overline{y}] = \int_{x_1}^{x_2} f(\overline{y}(x), \overline{y}'(x), x)\, dx \ . \tag{2.1.5}$$

Mas, uma vez que, por hipótese, a curva $y(x)$ fornece um extremo (mínimo, máximo ou ponto estacionário) de J, a função $\Phi(\epsilon)$ deve passar por um extremo para $\epsilon = 0$, pois,

[1] Suporemos que tal função existe. A questão da existência de soluções para problemas variacionais é um problema matemático muito delicado (Courant & Hilbert 1953).

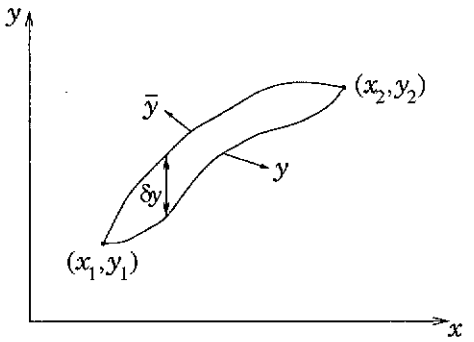


Figura 2.1: Variação de uma curva.

neste caso, $\overline{y}$ torna-se idêntica a y. Portanto, uma condição necessária para que $y(x)$ extremize J é

$$0 = \left(\frac{d\Phi}{d\epsilon}\right)_{\epsilon=0} = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial \overline{y}}\frac{\partial \overline{y}}{\partial \epsilon} + \frac{\partial f}{\partial \overline{y}'}\frac{\partial \overline{y}'}{\partial \epsilon}\right)_{\epsilon=0} dx \ , \tag{2.1.6}$$

decorrente da diferenciação de (2.1.5) sob o sinal de integral. Usando

$$\frac{\partial \overline{y}}{\partial \epsilon} = \eta \quad , \quad \frac{\partial \overline{y}'}{\partial \epsilon} = \eta' \ , \tag{2.1.7}$$

a Eq.(2.1.6) escreve-se (pois $\overline{y} = y$ para $\epsilon = 0$)

$$\left(\frac{d\Phi}{d\epsilon}\right)_{\epsilon=0} = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial y'}\eta'\right) dx = 0 \ . \tag{2.1.8}$$

Esta última equação é verdadeira com $\eta(x)$ arbitrária (exceto pela condição de anular-se nos limites de integração). A fim de extrair uma equação para $y(x)$, é necessário explorar a arbitrariedade de η. Mas para isto é preciso eliminar η' de (2.1.8), o que pode ser facilmente conseguido graças a uma integração por partes:

$$\int_{x_1}^{x_2} \frac{\partial f}{\partial y'}\eta' \, dx = \frac{\partial f}{\partial y'}\eta\Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} \eta \frac{d}{dx}\Big(\frac{\partial f}{\partial y'}\Big)dx = -\int_{x_1}^{x_2} \eta \frac{d}{dx}\Big(\frac{\partial f}{\partial y'}\Big)dx \tag{2.1.9}$$

onde usamos (2.1.4). Com a ajuda deste último resultado, a Eq.(2.1.8) reduz-se a

$$\int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\Big(\frac{\partial f}{\partial y'}\Big)\right]\eta dx = 0 \ . \tag{2.1.10}$$

Pode-se, agora, inferir uma equação diferencial para $y(x)$ apelando-se para um resultado importante.

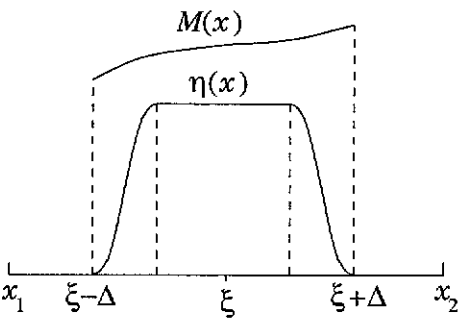


Figura 2.2: Lema fundamental do cálculo das variações.

Lema Fundamental do Cálculo das Variações. Se $M(x)$, $x_1 \leq x \leq x_2$, é uma função contínua tal que $\int_{x_1}^{x_2} M(x)\eta(x)dx = 0$ qualquer que seja a função contínua $\eta(x)$ com $\eta(x_1) = \eta(x_2) = 0$, então $M(x) \equiv 0$ em $[x_1, x_2]$.

Demonstração. Suponha $M(\xi) > 0$ para algum $x_1 < \xi < x_2$. Em virtude da continuidade de M, existe um subintervalo $(\xi - \Delta, \xi + \Delta)$ no qual $M(x) > c > 0$. Escolha a função contínua η (Fig. 2.2) de tal modo que $\eta = 1$ no intervalo $(\xi - \frac{\Delta}{2}, \xi + \frac{\Delta}{2})$, $\eta = 0$ fora do intervalo $(\xi - \Delta, \xi + \Delta)$ e $\eta(x) \geq 0$ para todo $x \in [x_1, x_2]$. A fim de dar à demonstração um caráter puramente analítico, eis aqui uma escolha possível para $\eta(x)$:

$$\eta(x) = \begin{cases} 0 & x_1 \leq x \leq \xi - \Delta \\ (4/\Delta^2)(x - \xi + \Delta)^2(1 + x - \xi + \Delta/2)^2 & \xi - \Delta \leq x \leq \xi - \Delta/2 \\ 1 & \xi - \Delta/2 \leq x \leq \xi + \Delta/2 \\ (4/\Delta^2)(x - \xi - \Delta)^2(1 + x - \xi - \Delta/2)^2 & \xi + \Delta/2 \leq x \leq \xi + \Delta \\ 0 & \xi + \Delta \leq x \leq x_2 \end{cases}$$

Claramente,

$$\int_{x_1}^{x_2} M(x)\eta(x)dx = \int_{\xi-\Delta}^{\xi+\Delta} M(x)\eta(x)dx \geq \int_{\xi-\Delta/2}^{\xi+\Delta/2} M(x)\eta(x)dx \geq c\Delta > 0 \ .$$

Esta contradição mostra que $M(\xi)$ não pode ser positivo, argumentos análogos aplicando-se na hipótese $M(\xi) < 0$. Como ξ é arbitrário, segue-se que $M(x) \equiv 0$ em $[x_1, x_2]$ e a demonstração está completa. □

Este lema aplicado à Eq.(2.1.10) conduz à *equação de Euler*[2]

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0 \ . \tag{2.1.11}$$

[2]Deduzida pela primeira vez por Leonhard Euler em 1744.

A equação diferencial de Euler é uma condição apenas *necessária*[3] para a existência de um extremo (máximo ou mínimo). Trata-se de uma equação diferencial de segunda ordem em cuja solução geral ocorrem duas constantes arbitrárias, o que normalmente permite satisfazer as condições de contorno de que a curva passe pelas extremidades fixas.

■ **Exemplo 2.1.2.** Determinar as geodésicas da esfera.
Solução. O comprimento de arco elementar em coordenadas esféricas é

$$ds^2 = dx^2 + dy^2 + dz^2 = dr^2 + r^2 d\theta^2 + r^2 \operatorname{sen}^2\theta \, d\varphi^2 \ .$$

Numa superfície esférica de raio R centrada na origem tem-se $r = R$, de modo que a distância entre dois pontos ao longo de uma curva $\varphi(\theta)$ sobre a superfície da esfera é dada por

$$s[\varphi] = R \int_{\theta_1}^{\theta_2} \sqrt{1 + \operatorname{sen}^2\theta \, \varphi'(\theta)^2} \, d\theta \ .$$

A curva geodésica é a que fornece a menor distância entre dois pontos, isto é, aquela que minimiza $s[\varphi]$. Conseqüentemente, $\varphi(\theta)$ tem que satisfazer a equação de Euler com $f(\varphi, \varphi', \theta) = \sqrt{1 + \operatorname{sen}^2\theta \, \varphi'(\theta)^2}$:

$$\frac{\partial f}{\partial \varphi} - \frac{d}{d\theta}\left(\frac{\partial f}{\partial \varphi'}\right) = 0 \Longrightarrow \frac{d}{d\theta}\left(\frac{\partial f}{\partial \varphi'}\right) = 0 \Longrightarrow \frac{\partial f}{\partial \varphi'} = \frac{\operatorname{sen}^2\theta \, \varphi'}{[1 + \operatorname{sen}^2\theta \, \varphi'(\theta)^2]^{1/2}} = C_1 \ ,$$

onde C_1 é uma constante arbitrária. Resolvendo esta última equação para φ' resulta

$$\frac{d\varphi}{d\theta} = \frac{C_1}{\operatorname{sen}\theta \sqrt{\operatorname{sen}^2\theta - C_1^2}} = \frac{C_1}{\operatorname{sen}^2\theta \sqrt{1 - C_1^2/\operatorname{sen}^2\theta}} = \frac{C_1/\operatorname{sen}^2\theta}{\sqrt{(1 - C_1^2) - C_1^2 \cot^2\theta}} \ .$$

Uma integração elementar fornece

$$\varphi = -\cos^{-1}\left(\frac{C_1 \cot\theta}{\sqrt{1 - C_1^2}}\right) + C_2 \ ,$$

donde

$$C \cot\theta = \cos(\varphi - C_2) \quad , \quad C = \frac{C_1}{\sqrt{1 - C_1^2}} \ .$$

Uma forma equivalente desta última equação é

$$C \cot\theta = A \cos\varphi + B \operatorname{sen}\varphi \quad , \quad A = \cos C_2 \ , \ B = \operatorname{sen} C_2 \ .$$

Para melhor compreender o que esta última equação representa, multipliquemo-la por $R \operatorname{sen}\theta$, obtendo como resultado

$$CR \cos\theta = AR \operatorname{sen}\theta \cos\varphi + BR \operatorname{sen}\theta \operatorname{sen}\varphi \ .$$

[3] São conhecidas condições suficientes de aplicação relativamente simples, pelo menos no caso de extremo fraco (Elsgoltz 1969; Gelfand & Fomin 1963).

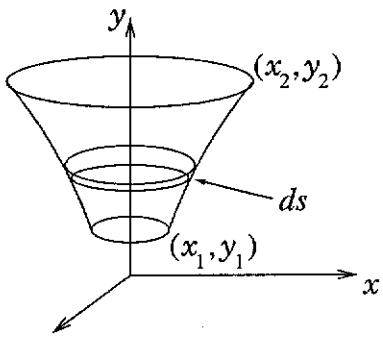


Figura 2.3: Superfície mínima de revolução.

Portanto, a curva geodésica é a interseção da superfície esférica $r = R$ com o plano passando pela origem caracterizado pela equação

$$Cz = Ax + By \ .$$

Assim, as geodésicas da superfície esférica são arcos de grandes círculos, isto é, círculos cujos centros coincidem com o centro da esfera. ■

■ **Exemplo 2.1.3.** Encontrar a superfície de revolução de menor área.

Solução. Considere uma superfície gerada pela rotação em torno do eixo y de uma curva plana passando por dois pontos fixos (Fig. 2.3) O problema consiste em encontrar a curva para a qual a área da superfície de revolução é mínima. A área de uma fita infinitesimal da superfície é $dA = 2\pi x ds = 2\pi x\sqrt{1+y'^2}\,dx$, de modo que a área total da superfície é

$$A[y] = 2\pi \int_{x_1}^{x_2} x\sqrt{1+y'^2}\,dx \ .$$

A equação de Euler deste problema com $f(y, y', x) = x\sqrt{1+y'^2}$ escreve-se

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0 \implies \frac{\partial f}{\partial y'} = a \implies \frac{xy'}{\sqrt{1+y'^2}} = a \ ,$$

onde a é uma constante de integração arbitrária (obviamente menor que x_1). Resolvendo esta última equação para y' obtém-se

$$y' = \frac{a}{\sqrt{x^2-a^2}} \ ,$$

donde, por uma integração imediata,

$$y = a\cosh^{-1}\left(\frac{x}{a}\right) + b$$

ou

$$x = a\cosh\frac{y-b}{a} \ .$$

A curva buscada é um arco de catenária. As constantes a e b devem ser determinadas exigindo que a catenária passe pelos pontos extremos. Para certas posições dos pontos extremos podem existir duas catenárias distintas, apenas uma delas fornecendo o mínimo absoluto, ou o verdadeiro mínimo ser alcançado numa curva que não pertence à classe das catenárias e nem sequer possui derivada contínua (Weinstock 1952; Arfken 1985). ■

■ **Exemplo 2.1.4.** Resolver o problema da braquistócrona.

Solução. O problema inaugural e mais célebre do cálculo das variações é o da linha de deslizamento mais rápido, ou *braquistócrona*,[4] proposto como desafio aos matemáticos europeus, em 1696, por Johann Bernoulli e resolvido, muito antes do aparecimento da equação de Euler, pelo próprio desafiador e, independentemente, por l'Hôpital, Jakob Bernoulli, Gottfried Wilhelm Leibnitz e Isaac Newton.[5] O problema consiste em determinar a curva unindo dois pontos dados P e Q, não pertencentes a uma mesma reta vertical, que possua a propriedade de que, sob a ação da gravidade, uma partícula deslize (sem atrito) ao longo dela no menor tempo possível. Escolhamos a origem no ponto P e orientemos o eixo y verticalmente para baixo. Depois de cair uma altura y o módulo da velocidade da partícula é $ds/dt = v = \sqrt{2gy}$. Usando $ds = \sqrt{1+y'^2}dx$, o tempo gasto para deslizar da posição P= $(0,0)$ até Q= (x_0, y_0) é

$$T[y] = \int_P^Q \frac{ds}{v} = \frac{1}{\sqrt{2g}} \int_0^{x_0} \left(\frac{1+y'^2}{y}\right)^{1/2} dx \quad , \quad y(0) = 0 \quad , \quad y(x_0) = y_0 \ .$$

A resolução deste problema é facilitada pela escolha de y como variável independente e da curva na forma $x(y)$. Neste caso, com $x' = dx/dy$,

$$T[x] = \frac{1}{\sqrt{2g}} \int_0^{y_0} \left(\frac{1+x'^2}{y}\right)^{1/2} dy \quad , \quad x(0) = 0 \quad , \quad x(y_0) = x_0 \ .$$

Para o problema assim formulado tem-se $f(x, x', y) = \sqrt{(1+x'^2)/y}$ e equação de Euler permite uma primeira integração imediata:

$$\frac{\partial f}{\partial x} - \frac{d}{dy}\left(\frac{\partial f}{\partial x'}\right) = 0 \quad \Longrightarrow \quad \frac{\partial f}{\partial x'} = \frac{x'}{\sqrt{y(1+x'^2)}} = C_1 \ ,$$

onde C_1 é uma constante arbitrária. Esta última equação integra-se mais facilmente introduzindo-se o parâmetro t através da equação $x' = \tan t$. Então, com $C_2 = 1/C_1^2$,

$$y = C_2 \frac{\tan^2 t}{1+\tan^2 t} = C_2 \operatorname{sen}^2 t = \frac{C_2}{2}(1 - \cos 2t) \ ;$$

$$dx = \tan t \, dy = 2C_2 \operatorname{sen}^2 t \, dt = C_2(1 - \cos 2t)\, dt \ \Longrightarrow \ x = \frac{C_2}{2}(2t - \operatorname{sen} 2t) + C_3 \ .$$

Tendo em conta que $C_3 = 0$ porque $y = 0$ para $x = 0$, e fazendo $\theta = 2t$, resultam as equações paramétricas de uma família de ciclóides

$$x = C(\theta - \operatorname{sen}\theta) \quad , \quad y = C(1 - \cos\theta) \ .$$

[4] Do grego *brachistos* (brevíssimo) e *chronos* (tempo).

[5] Na noite de 29 de janeiro de 1697, quando recebeu a carta-desafio, Newton não dormiu até resolver o problema, o que se deu por volta de quatro horas da manhã. Em seguida, a solução foi remetida anonimamente para Bernoulli. Ao ler a solução chegada da Inglaterra, Johann Bernoulli, segundo suas próprias palavras, reconheceu imediatamente o seu autor "como se reconhece o leão por sua pata".

A constante $C = C_2/2$ é determinada pela condição de que a curva passe pelo ponto Q= (x_0, y_0). Portanto, a braquistócrona é uma arco de cicloide. ■

2.2 Notação Variacional

A variação δy, representada na Fig. 2.1, é definida por $\delta y = \epsilon\eta$, de modo que a Eq.(2.1.3) escreve-se

$$\overline{y} = y + \delta y \ . \tag{2.2.1}$$

Analogamente, a variação do funcional J é definida por

$$\delta J = \epsilon \left(\frac{d\Phi}{d\epsilon}\right)_{\epsilon=0} = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial y'}\delta y'\right) dx \ , \tag{2.2.2}$$

onde usamos a primeira igualdade em (2.1.8) e

$$\delta y' = \epsilon\eta' = (\epsilon\eta)' = (\delta y)' \equiv \frac{d}{dx}(\delta y) \ . \tag{2.2.3}$$

A notação aqui adotada é análoga à notação do cálculo diferencial, no qual a expressão $\epsilon f'(x) \equiv df$ com parâmetro arbitrário ϵ é chamada de diferencial da função $f(x)$. Uma integração por partes, com o uso de $\delta y(x_1) = \delta y(x_2) = 0$, reduz a Eq.(2.2.2) à forma

$$\delta J = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right] \delta y \, dx \ . \tag{2.2.4}$$

Assim, uma condição necessária para um extremo de um funcional $J[y]$ é que sua variação δJ anule-se para δy arbitrária, exceto pela condição de anular-se nas extremidades do intervalo de integração. Em geral, as funções $y(x)$ para as quais $\delta J = 0$ são chamadas de *extremantes* ou *curvas estacionárias*. Esta terminologia se justifica porque $\delta J = 0$ significa apenas que uma variação δy da curva y deixa o funcional $J[y]$ estacionário, não havendo qualquer garantia de que a curva forneça um máximo ou mínimo do funcional.

2.3 Princípio de Hamilton e Equações de Lagrange

A simples mudança de notação

$$x \to t \ , \quad y \to q \ , \quad y' \equiv \frac{dy}{dx} \to \dot{q} \equiv \frac{dq}{dt} \ , \quad f \to L \ , \quad J \to S \tag{2.3.1}$$

mostra que a equação de Euler reduz-se à equação de Lagrange

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = 0 \ , \tag{2.3.2}$$

a qual, portanto, decorre do princípio variacional

$$\delta S \equiv \delta \int_{t_1}^{t_2} L(q, \dot{q}, t)\, dt = 0 \tag{2.3.3}$$

com $\delta q(t_1) = \delta q(t_2) = 0$. A generalização para sistemas holônomos com um número qualquer de graus de liberdade é imediata. Seja S a *ação* definida por

$$S = \int_{t_1}^{t_2} L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t)\, dt \tag{2.3.4}$$

e considere

$$\begin{aligned} \overline{q}_1(t) &= q_1(t) + \delta q_1(t) \ , \\ &\vdots \qquad\quad \vdots \qquad\quad \vdots \\ \overline{q}_n(t) &= q_n(t) + \delta q_n(t) \ , \end{aligned} \tag{2.3.5}$$

com as variações $\delta q_1, \ldots, \delta q_n$ independentes entre si e arbitrárias, exceto pelas condições nos extremos $\delta q_k(t_1) = \delta q_k(t_2) = 0$, $k = 1, \ldots, n$. Destaque-se que é a mútua independência das coordenadas generalizadas que assegura que cada uma delas pode ser variada independentemente das demais.

A variação da ação $S = \int_{t_1}^{t_2} L dt$ é dada por

$$\delta S = \int_{t_1}^{t_2} \sum_k \left(\frac{\partial L}{\partial q_k} \delta q_k + \frac{\partial L}{\partial \dot{q}_k} \delta \dot{q}_k \right) dt \ . \tag{2.3.6}$$

Uma integração por partes fornece

$$\delta S = \int_{t_1}^{t_2} dt \sum_k \left[\frac{\partial L}{\partial q_k} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k} \right) \right] \delta q_k + \sum_k \frac{\partial L}{\partial \dot{q}_k} \delta q_k \Bigg|_{t_1}^{t_2} \ . \tag{2.3.7}$$

Levando em conta que as variações das coordenadas generalizadas anulam-se nos pontos extremos, esta última equação reduz-se a

$$\delta S = \int_{t_1}^{t_2} dt \sum_k \left[\frac{\partial L}{\partial q_k} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k} \right) \right] \delta q_k \ . \tag{2.3.8}$$

Impondo $\delta S = 0$ somos conduzidos a

$$\int_{t_1}^{t_2} dt \sum_k \left[\frac{\partial L}{\partial q_k} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k} \right) \right] \delta q_k = 0 \ . \tag{2.3.9}$$

Uma vez que os δqs são mutuamente independentes, podemos tomar todos iguais a zero exceto um particular δq_{k_0}. Neste caso, a soma em (2.3.9) reduz-se a um único termo correspondente a $k = k_0$. Mas, como δq_{k_0} é uma função arbitrária, o lema fundamental do cálculo das variações estabelece que o coeficiente de δq_{k_0} em (2.3.9) é identicamente

nulo. Finalmente, como o argumento anterior é aplicável a qualquer k_0, conclui-se que (2.3.9) equivale a

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = 0 \quad , \quad k = 1, \ldots, n \ . \tag{2.3.10}$$

Isto completa a dedução das equações de Lagrange a partir do princípio variacional $\delta S = 0$ para sistemas holônomos.

A ocasião é propícia para introduzir uma sugestiva descrição geométrica da dinâmica lagrangiana. Como já vimos, o espaço cartesiano de n dimensões cujos pontos são as n-uplas $(q_1, \ldots, q_n)$ é conhecido como espaço de configuração. O nome é adequado: cada ponto do espaço de configuração corresponde a um conjunto de valores determinados de $(q_1, \ldots, q_n)$ que define univocamente a configuração do sistema mecânico no instante considerado. À medida que o tempo passa, o estado do sistema se modifica e o ponto representativo do sistema descreve uma curva no espaço de configuração, já que as equações $q_1 = q_1(t), \ldots, q_n = q_n(t)$ são a representação paramétrica de uma curva tendo t como parâmetro. Essa trajetória representa o movimento do sistema em sua totalidade no espaço de configuração, e este, dada a sua sua natureza abstrata, não costuma ter qualquer relação com o espaço físico tridimensional. As coordenadas generalizadas não são necessariamente coordenadas de posição e a trajetória do sistema no espaço de configuração não guarda, em geral, qualquer semelhança com a curva descrita no espaço tridimensional por qualquer partícula constituinte do sistema. De agora em diante, salvo menção em contrário, designaremos simplesmente por q a n-upla $(q_1, \ldots, q_n)$, isto é, $q \equiv (q_1, \ldots, q_n)$. A evolução dinâmica corresponde a uma trajetória traçada no espaço de configuração e podemos enunciar o seguinte princípio fundamental.

Princípio de Hamilton. Dado um sistema mecânico holônomo descrito pela lagrangiana $L(q, \dot{q}, t)$, seu movimento do instante t_1 ao instante t_2 é tal que a ação

$$S = \int_{t_1}^{t_2} L(q, \dot{q}, t)\, dt \tag{2.3.11}$$

é mínima (mais geralmente, estacionária) para a trajetória real, mantidos fixos os pontos inicial e final da trajetória no espaço de configuração.

Lagrangianas Equivalentes

O princípio de Hamilton, que também costuma ser chamado de *princípio da mínima ação*,[6] tem como conseqüência importante que as mesmas equações de movimento são geradas por duas lagrangianas que só diferem pela derivada total em relação ao tempo de uma função arbitrária das coordenadas generalizadas e do tempo.

Definição 2.3.1. Duas lagrangianas $\bar{L}(q, \dot{q}, t)$ e $L(q, \dot{q}, t)$ são ditas *equivalentes* se elas diferem pela derivada total em relação ao tempo de uma função arbitrária das coordenadas generalizadas e do tempo $f(q, t)$:

$$\bar{L}(q, \dot{q}, t) = L(q, \dot{q}, t) + \frac{d}{dt} f(q, t) \ . \tag{2.3.12}$$

[6] Uma discussão encantadora e envolvente do princípio da mínima ação encontra-se em Feynman, Leighton & Sands (1963), volume II, capítulo 19.

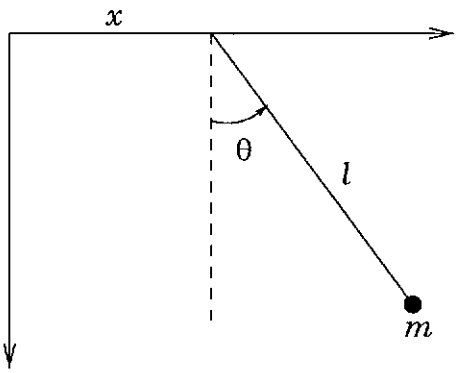


Figura 2.4: Pêndulo plano com ponto de suspensão móvel.

Teorema 2.3.1. Lagrangianas equivalentes dão lugar às mesmas equações de movimento.

Demonstração. A ação $\bar{S}$ associada a $\bar{L}$ é dada por

$$\bar{S} = \int_{t_1}^{t_2} \bar{L}(q, \dot{q}, t)\, dt = \int_{t_1}^{t_2} L(q, \dot{q}, t)\, dt + \int_{t_1}^{t_2} \frac{df}{dt}\, dt = S + f(q(t_2), t_2) - f(q(t_1), t_1) \ .$$

Mas, como a variação da ação mantém os extremos $q(t_1)$ e $q(t_2)$ fixos, $\delta\bar{S} = \delta S$. Portanto, as condições $\delta\bar{S} = 0$ e $\delta S = 0$ são idênticas, demonstrando que $\bar{L}$ e L produzem exatamente as mesmas equações de movimento. □

■ **Exercício 2.3.1.** Demonstre, por substituição direta nas equações de Lagrange, que lagrangianas equivalentes geram as mesmas equações de movimento. ■

■ **Exemplo 2.3.1.** Discutir a lagrangiana para um pêndulo plano cujo ponto de suspensão desloca-se sobre uma reta horizontal com velocidade constante.

Solução. Usando as coordenadas indicadas na Fig. 2.4, temos

$$T = \frac{m}{2}\left[(\dot{x} + l\dot{\theta}\cos\theta)^2 + (-l\dot{\theta}\,\mathrm{sen}\,\theta)^2 \right] \ ,$$

donde

$$L = T - V = \frac{m}{2}(\dot{x}^2 + l^2\dot{\theta}^2 + 2l\dot{x}\dot{\theta}\cos\theta) + mgl\cos\theta \ .$$

Se $\dot{x} = v =$ constante podemos escrever

$$L = \frac{ml^2}{2}\dot{\theta}^2 + mgl\cos\theta + \frac{d}{dt}\left(\frac{m}{2}v^2 t + mlv\,\mathrm{sen}\,\theta\right) \ .$$

Descartando a derivada total, a lagrangiana resultante coincide com a que seria obtida por um

observador em relação ao qual o ponto de suspensão do pêndulo é fixo. Você é capaz de explicar por quê? ■

É possível demonstrar (Problema 2.6) que duas lagrangianas só produzem *exatamente* as mesmas equações de movimento se diferem pela derivada total em relação ao tempo de uma função das coordenadas generalizadas e do tempo. No entanto, há lagrangianas que não diferem por uma derivada total mas geram equações de movimento *equivalentes*, isto é, equações que não são idênticas mas possuem as mesmas soluções (Problema 2.7).

2.4 Princípio de Hamilton no Caso Não-Holônomo

Na seção anterior admitimos que o sistema mecânico era descrito por coordenadas independentes entre si, e tal independência foi crucial para permitir a dedução de (2.3.10) a partir de (2.3.9). Tais coordenadas generalizadas mutuamente independentes sempre existem quando todos os vínculos a que o sistema está sujeito são holônomos. Quando há vínculos não-holônomos presentes, é em geral impossível introduzir coordenadas generalizadas de tal modo que as equações de vínculo sejam identicamente satisfeitas. É possível, ainda assim, deduzir as equações de movimento a partir do princípio de Hamilton no caso especial em que os vínculos não-holônomos são equações diferenciais da forma

$$\sum_{k=1}^{n} a_{lk}dq_k + a_{lt}dt = 0 \quad , \quad l = 1, \ldots, p \ , \tag{2.4.1}$$

cujos coeficientes a_{lk} e a_{lt} são funções somente de $q_1, \ldots, q_n$, e t. Suponhamos, portanto, que o sistema seja descrito por n coordenadas $(q_1, \ldots, q_n)$ e esteja submetido aos p vínculos diferenciais independentes (2.4.1), o índice l servindo para distinguir as equações de vínculo umas das outras. Apesar da forma aparentemente restrita, os vínculos do tipo (2.4.1) abrangem quase todos os casos de vínculos não-holônomos de interesse físico.

Seja L a lagrangiana do sistema *escrita como se não houvesse vínculos*. O potencial generalizado em L refere-se apenas às forças aplicadas, de modo que as forças de vínculo responsáveis pela validade das Eqs.(2.4.1) não estão incluídas na lagrangiana. O princípio de Hamilton $\delta S = 0$ implica

$$\int_{t_1}^{t_2} dt \sum_k \left[\frac{\partial L}{\partial q_k} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) \right] \delta q_k = 0 \ , \tag{2.4.2}$$

mas, agora, *não* podemos inferir que o coeficiente da cada δq_k é zero porque os δqs *não são mutuamente independentes*. Com efeito, as Eqs.(2.3.5) mostram que cada δq_k é um deslocamento virtual, pois o tempo permanece fixo quando se executa a variação que leva $q_k(t)$ em $\bar{q}_k(t) = q_k(t) + \delta q_k(t)$. Mas, conforme as Eqs.(2.4.1), os deslocamentos virtuais devem obedecer a

$$\sum_{k=1}^{n} a_{lk}\delta q_k = 0 \quad , \quad l = 1, \ldots, p \tag{2.4.3}$$

para que haja compatibilidade com os vínculos (usamos $dt = 0$ para deslocamentos virtuais). As n variações $\delta q_1, \ldots, \delta q_n$ têm que satisfazer as p equações (2.4.3), de modo

que apenas $(n-p)$ variações dos qs são independentes entre si.[7] Trata-se, portanto, da determinação de um extremo condicionado para o funcional S, e o tratamento segue a linha do método dos multiplicadores de Lagrange do cálculo diferencial.

Multiplicadores de Lagrange

A equação

$$\int_{t_1}^{t_2} dt \sum_{l=1}^{p} \lambda_l \Big(\sum_{k=1}^{n} a_{lk} \delta q_k \Big) = \int_{t_1}^{t_2} dt \sum_{k=1}^{n} \Big(\sum_{l=1}^{p} \lambda_l a_{lk} \Big) \delta q_k = 0 \tag{2.4.4}$$

é uma conseqüência óbvia de (2.4.3) para valores arbitrários dos multiplicadores de Lagrange $\lambda_1(q,\dot{q},t),\ldots,\lambda_n(q,\dot{q},t)$. Adicionando (2.4.2) a (2.4.4) resulta

$$\int_{t_1}^{t_2} dt \sum_{k=1}^{n} \left\{ \frac{\partial L}{\partial q_k} - \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_k}\Big) + \sum_{l=1}^{p} \lambda_l a_{lk} \right\} \delta q_k = 0 \ . \tag{2.4.5}$$

Como os δqs não são independentes, nada podemos afirmar a respeito do coeficiente de cada δq_k nesta última equação. Com uma numeração adequada das variáveis, podemos escolher as $(n-p)$ primeiras variações $\delta q_1,\ldots,\delta q_{n-p}$ como mutuamente independentes, as p últimas variações sendo determinadas em termos das $(n-p)$ primeiras pela resolução das p equações (2.4.3). Por outro lado, temos p multiplicadores de Lagrange à nossa disposição e podemos escolhê-los de tal modo que os coeficientes dos p últimos δqs em (2.4.5) sejam nulos, isto é,

$$\frac{\partial L}{\partial q_k} - \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_k}\Big) + \sum_{l=1}^{p} \lambda_l a_{lk} = 0 \quad , \quad k = n-p+1,\ldots,n \ . \tag{2.4.6}$$

Com os λs determinados pelas equações (2.4.6), a Eq.(2.4.5) reduz-se a

$$\int_{t_1}^{t_2} dt \sum_{k=1}^{n-p} \left\{ \frac{\partial L}{\partial q_k} - \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_k}\Big) + \sum_{l=1}^{p} \lambda_l a_{lk} \right\} \delta q_k = 0 \ , \tag{2.4.7}$$

que só envolve os $(n-p)$ primeiros δqs independentes entre si, implicando

$$\frac{\partial L}{\partial q_k} - \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_k}\Big) + \sum_{l=1}^{p} \lambda_l a_{lk} = 0 \quad , \quad k = 1,\ldots,n-p \ . \tag{2.4.8}$$

Em síntese, as Eqs.(2.4.6) e (2.4.8) mostram que as equações de movimento do sistema são

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_k}\Big) - \frac{\partial L}{\partial q_k} = \sum_{l=1}^{p} \lambda_l a_{lk} \quad , \quad k = 1,\ldots,n \ . \tag{2.4.9}$$

Mas isto é insuficiente, pois agora temos apenas n equações para $(n+p)$ incógnitas,

[7] Os sistemas holônomos distinguem-se dos não-holônomos pela propriedade de o número de deslocamentos virtuais independentes coincidir com o número mínimo de coordenadas necessárias para especificar a configuração do sistema.

a saber, as n coordenadas q_k e os p multiplicadores de Lagrange λ_l. As p equações adicionais são, evidentemente, as equações de vínculo (2.4.1), que podem ser escritas na forma equivalente

$$\sum_{k=1}^{n} a_{lk}\dot{q}_k + a_{lt} = 0 \quad , \quad l = 1, \ldots, p \ . \tag{2.4.10}$$

As equações (2.4.9) e (2.4.10) constituem um conjunto de $(n+p)$ equações para $(n+p)$ incógnitas, e permitem a determinação unívoca do movimento do sistema.

Forças de Vínculo

Resta investigar o significado físico dos multiplicadores de Lagrange, que também são determinados no processo de resolução das equações de movimento. Imaginemos os vínculos removidos e forças generalizadas Q'_k atuando sem provocar nenhuma alteração no movimento do sistema. Ora, isto só ocorreria se Q'_k fossem as forças de vínculo, pois somente assim o sistema obedeceria às restrições (2.4.10). Mas, levando em conta que L só inclui as forças aplicadas, as forças adicionais Q'_k apareceriam nas equações de movimento na forma (1.6.13), isto é,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = Q'_k \ . \tag{2.4.11}$$

Para que os movimentos sejam idênticos, as equações de movimento (2.4.9) e (2.4.11) precisam ser idênticas. Em outras palavras,

$$Q'_k = \sum_{l=1}^{p} \lambda_l a_{lk} \tag{2.4.12}$$

é a k-ésima força generalizada de vínculo. Assim, além do movimento do sistema, na presente formulação as forças de vínculo aparecem como parte da resposta. É preciso não esquecer que as forças generalizadas de vínculo estão relacionadas com as forças de vínculo propriamente ditas por intermédio de (1.4.6).

A conclusão da análise acima é que, desde que adequadamente modificadas, as equações de Lagrange são válidas mesmo na presença de vínculos não-holônomos. Além do movimento do sistema mecânico, as forças de vínculo são determinadas como importante subproduto do formalismo dos multiplicadores de Lagrange.[8]

■ **Exercício 2.4.1.** Prove que o trabalho virtual total das forças generalizadas de vínculo (2.4.12) é zero. Isto mostra que mesmo no caso de vínculos não-holônomos o princípio de Hamilton está em harmonia com o princípio de d'Alembert. ■

Comentário. No caso mais geral em que as equações de vínculo são da forma

$$g_l(q, \dot{q}, t) = 0 \quad , \quad l = 1, \ldots, p \ , \tag{2.4.13}$$

[8] É possível formular as equações de movimento para sistemas não-holônomos sem lançar mão de multiplicadores de Lagrange. Trata-se de um método baseado na introdução das chamadas quase-coordenadas e que culmina nas equações de Gibbs-Appell (Pars 1965; Gantmacher 1970; Desloge 1982).

o resultado considerado correto (Rund 1966; Saletan & Cromer 1970) para as equações de movimento é

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = \sum_{l=1}^{p} \lambda_l \frac{\partial g_l}{\partial \dot{q}_k} \quad , \qquad k = 1, \ldots, n \ . \tag{2.4.14}$$

É claro que estas equações reduzem-se a (2.4.9) quando as equações de vínculo envolvem as velocidades linearmente.

■ **Exemplo 2.4.1.** Como primeira ilustração puramente acadêmica, considere uma partícula movendo-se livremente num plano exceto pelo vínculo não-holônomo

$$\dot{x} - \omega y = 0 \ , \tag{2.4.15}$$

onde ω é uma constante. Empregando a notação $q_1 = x$ e $q_2 = y$, vemos que (2.4.15) é da forma (2.4.10) com $a_{11} = 1$, $a_{12} = 0$, $a_{1t} = -\omega y$. Como há apenas um vínculo, l toma somente o valor $l = 1$. Usando-se a lagrangiana

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) \tag{2.4.16}$$

as Eqs.(2.4.9) adquirem a forma

$$m\ddot{x} = \lambda_1 a_{11} = \lambda_1 \quad , \quad m\ddot{y} = \lambda_1 a_{12} = 0 \ . \tag{2.4.17}$$

A solução geral da segunda das Eqs.(2.4.17) é

$$y = y_0 + v_0 t \ , \tag{2.4.18}$$

com y_0 e v_0 constantes arbitrárias. Levando (2.4.18) em (2.4.15) e integrando, obtém-se

$$x = x_0 + \omega y_0 t + \frac{1}{2}\omega v_0 t^2 \ , \tag{2.4.19}$$

com x_0 constante arbitrária. A solução geral para o movimento do sistema está contida em (2.4.18) e (2.4.19). Finalmente, o multiplicador de Lagrange é determinado por

$$\lambda_1 = m\ddot{x} = m\omega\dot{y} = m\omega v_0 \ ,$$

resultando

$$Q'_1 \equiv Q'_x \equiv F'_x = \lambda_1 a_{11} = \lambda_1 = m\omega\dot{y} \quad , \quad F'_y = \lambda_1 a_{12} = 0 \tag{2.4.20}$$

para as componentes cartesianas da força de vínculo. ■

■ **Exercício 2.4.2.** Prove que o vínculo (2.4.15) não é holônomo. ■

■ **Exercício 2.4.3.** Resolva as equações newtonianas de movimento para uma partícula numa plano sujeita à força $\mathbf{F}'$ com as componentes (2.4.20). A solução geral obtida satisfaz

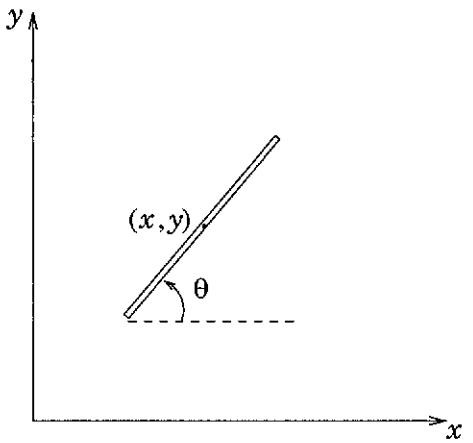


Figura 2.5: Patinete num plano horizontal.

automaticamente o vínculo (2.4.15)? Sugestão: o que deve acontecer com as condições iniciais para que os vínculos sejam respeitados durante todo o movimento do sistema? ■

Advertência. Vínculos não-integráveis *não podem* ser substituídos na lagrangiana para eliminar variáveis. Se uma variável for assim eliminada da lagrangiana, as configurações acessíveis estarão sendo indevidamente reduzidas, pois vínculos não-integráveis restrigem apenas as velocidades. Por exemplo, se a variável x for completamente eliminada da lagrangiana (2.4.16) por meio de (2.4.15), a nova lagrangiana $\bar{L}(y,\dot{y})$ assim obtida conduzirá a uma equação de movimento para y cuja solução é completamente diferente de (2.4.18), que é a solução correta (verifique você mesmo). Em muitos casos a eliminação de variáveis pela substituição de vínculos não-integráveis na lagrangiana é possível, mas as equações de movimento corretas são as equações de Voronec (Neĭmark & Fufaev 1972; Lemos 2003).

■ **Exemplo 2.4.2.** Usando o método dos multiplicadores de Lagrange, descrever o movimento de um patinete num plano horizontal.

Solução. Como modelo supersimplificado de patinete, tomemos uma haste rígida, delgada, homogênea de comprimento l restrita a mover-se de tal modo que a velocidade do centro de massa seja sempre paralela à haste (Fig. 2.5). Como coordenadas utilizaremos o vetor posição do centro de massa $\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}}$ e o ângulo θ que a haste faz com o eixo x. Sendo $\hat{\mathbf{n}} = \cos\theta\,\hat{\mathbf{i}} + \operatorname{sen}\theta\,\hat{\mathbf{j}}$ o vetor unitário ao longo da haste, o vínculo exprime-se pela equação $\dot{\mathbf{r}} \times \hat{\mathbf{n}} = 0$ ou, em componentes,

$$\dot{x}\operatorname{sen}\theta - \dot{y}\cos\theta = 0 \ . \tag{2.4.21}$$

Utilizando a notação $q_1 = x$, $q_2 = y$, $q_3 = \theta$, identificamos $a_{11} = \operatorname{sen}\theta$, $a_{12} = -\cos\theta$, $a_{13} = a_{1t} = 0$ por simples comparação entre (2.4.21) e (2.4.10). Dado que $V = 0$, a lagrangiana é simplesmente a energia cinética, a qual, pela Eq.(1.1.22), é a soma da energia cinética do centro de massa com a energia de rotação da haste em torno do centro de massa:

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) + \frac{I}{2}\dot{\theta}^2 \ , \tag{2.4.22}$$

onde I é o momento de inércia da haste relativamente a um eixo perpendicular a ela passando pelo centro de massa. As equações de movimento são

$$m\ddot{x} = \lambda_1 \operatorname{sen}\theta \quad , \quad m\ddot{y} = -\lambda_1 \cos\theta \quad , \quad I\ddot{\theta} = 0 \ . \tag{2.4.23}$$

A equação para θ tem por solução

$$\theta = \theta_0 + \Omega t \ , \tag{2.4.24}$$

com θ_0 e Ω constantes arbitrárias. Combinando as duas primeiras equações de movimento deduzimos

$$\ddot{x}\cos\theta + \ddot{y}\operatorname{sen}\theta = 0 \ . \tag{2.4.25}$$

Da equação de vínculo (2.4.21) decorre

$$\dot{y} = \dot{x}\tan\theta \implies \ddot{y} = \ddot{x}\tan\theta + \Omega\dot{x}\sec^2\theta \ . \tag{2.4.26}$$

Introduzindo este resultado em (2.4.25) resulta

$$\frac{\ddot{x}}{\cos\theta} + \Omega\dot{x}\frac{\operatorname{sen}\theta}{\cos^2\theta} = 0 \implies \frac{d}{dt}\left(\frac{\dot{x}}{\cos\theta}\right) = 0 \implies \dot{x} = C'\cos(\Omega t + \theta_0) \ . \tag{2.4.27}$$

Uma nova integração fornece

$$x = x_0 + C\operatorname{sen}(\Omega t + \theta_0) \ , \tag{2.4.28}$$

onde x_0 e $C\,(= C'/\Omega)$ são constantes arbitrárias. A inserção de (2.4.24) e (2.4.28) na primeira das Eqs.(2.4.26), seguida de uma integração, redunda em

$$y = y_0 - C\cos(\Omega t + \theta_0) \ . \tag{2.4.29}$$

O patinete gira em torno de seu próprio centro de massa com velocidade angular constante Ω, ao mesmo tempo que o centro da massa descreve uma circunferência no plano do movimento com a mesma velocidade angular Ω. Se $\Omega = 0$, o centro de massa do patinete executa um movimento retilíneo uniforme. ■

Vínculos Holônomos e Multiplicadores de Lagrange

Embora as equações da forma (2.4.10) não incluam os tipos mais gerais de vínculos não-holônomos, elas dão conta de vínculos holônomos. Com efeito, vínculos holônomos são da forma

$$f_l(q,t) = 0 \quad , \quad l = 1,\ldots,p \ , \tag{2.4.30}$$

ou, tomando a derivada total em relação ao tempo,

$$\sum_{k=1}^{n} \frac{\partial f_l}{\partial q_k}\dot{q}_k + \frac{\partial f_l}{\partial t} = 0 \quad , \quad l = 1,\ldots,p \ , \tag{2.4.31}$$

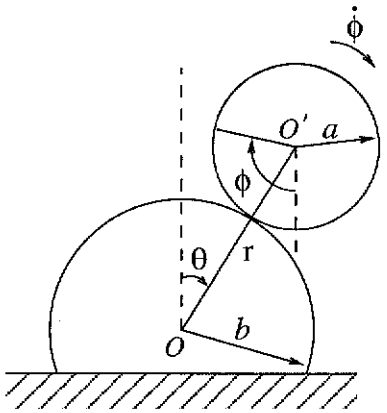


Figura 2.6: Cilindro rolando sem deslizar sobre cilindro fixo.

que são da forma (2.4.10) com

$$a_{lk} = \frac{\partial f_l}{\partial q_k} \quad , \quad a_{lt} = \frac{\partial f_l}{\partial t} \ . \tag{2.4.32}$$

Assim, o método dos multiplicadores de Lagrange pode ser aplicado a sistemas holônomos quando é inconveniente substituir os qs por um conjunto menor de variáveis independentes ou quando se deseja obter as forças de vínculo. À guisa de ilustração, apliquemos o formalismo dos multiplicadores de Lagrange a um caso holônomo.

■ **Exemplo 2.4.3.** Lançando mão da técnica dos multiplicadores de Lagrange, analisar o rolamento sem deslizamento de um cilindro sobre outro cilindro fixo.

Solução. Usaremos como coordenadas os ângulos θ e ϕ representados na Fig. 2.6, onde ϕ é o ângulo de rotação do cilindro móvel em torno do seu eixo de simetria, além de r, que é a distância entre os eixos dos cilindros. Note que r e θ são coordenadas polares do centro de massa do cilindro móvel relativamente ao centro de massa do cilindro fixo. A lagrangiana do problema deve ser escrita como se r, θ, ϕ fossem coordenadas independentes. A energia cinética do cilindro móvel compõe-se da energia cinética do centro de massa mais a energia de rotação em torno do centro de massa, isto é,

$$T = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{1}{2}I\dot{\phi}^2 \ , \tag{2.4.33}$$

onde $I = ma^2/2$ é o momento de inércia do cilindro em torno do seu eixo de simetria. A lagrangiana, portanto, é

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{ma^2}{4}\dot{\phi}^2 - mgr\cos\theta \ . \tag{2.4.34}$$

Neste problema há dois vínculos holônomos, a saber:

$$r\dot{\theta} = a\dot{\phi} \quad \text{(vínculo de rolar sem deslizar)} \ ; \tag{2.4.35}$$

$$r = a + b \quad \text{(vínculo de contato)} \ . \tag{2.4.36}$$

O vínculo de rolar sem deslizar exprime a exigência de que o ponto O' desloque-se com velocidade $v_\theta = r\dot{\theta}$ igual àquela gerada por sua rotação com velocidade angular $\dot{\phi}$ em torno do eixo instantâneo de rotação, que passa pelo ponto de contato entre os cilindros. Em forma diferencial, o vínculo de contato escreve-se

$$\dot{r} = 0 \ . \tag{2.4.37}$$

Introduzindo a notação $q_1 = r$, $q_2 = \theta$, $q_3 = \phi$, identificamos

$$a_{11} = 0 \ , \quad a_{12} = r \ , \quad a_{13} = -a \ , \quad a_{21} = 1 \ , \quad a_{22} = a_{23} = 0 \ , \tag{2.4.38}$$

onde $l = 1$ corresponde ao vínculo (2.4.35) e $l = 2$ ao vínculo (2.4.37).

As equações de movimento assumem a forma

$$m\ddot{r} - mr\dot{\theta}^2 + mg\cos\theta = \lambda_1 a_{11} + \lambda_2 a_{21} = \lambda_2 \ , \tag{2.4.39}$$

$$m\frac{d}{dt}(r^2\dot{\theta}) - mgr\,\mathrm{sen}\,\theta = \lambda_1 a_{12} + \lambda_2 a_{22} = r\lambda_1 \ , \tag{2.4.40}$$

$$\frac{ma^2}{2}\ddot{\phi} = \lambda_1 a_{13} + \lambda_2 a_{23} = -a\lambda_1 \ . \tag{2.4.41}$$

Com o emprego das equações de vínculo, as equações de movimento reduzem-se a

$$-m(a+b)\dot{\theta}^2 + mg\cos\theta = \lambda_2 \ , \tag{2.4.42}$$

$$m(a+b)^2\ddot{\theta} - mg(a+b)\,\mathrm{sen}\,\theta = (a+b)\lambda_1 \ , \tag{2.4.43}$$

$$\frac{m}{2}(a+b)\ddot{\theta} = -\lambda_1 \ . \tag{2.4.44}$$

Combinando as duas últimas equações de modo a eliminar λ_1, obtém-se

$$\frac{3}{2}(a+b)\ddot{\theta} - g\,\mathrm{sen}\,\theta = 0 \ . \tag{2.4.45}$$

A força de reação normal do cilindro fixo sobre o móvel é a componente radial da força de vínculo sobre o cilindro móvel, isto é,

$$N = Q'_r \equiv Q'_1 = \sum_{l=1}^{2} \lambda_l a_{l1} = \lambda_2 = mg\cos\theta - m(a+b)\dot{\theta}^2 \ . \tag{2.4.46}$$

É possível, agora, responder à seguinte pergunta: em que ponto o cilindro móvel se desprende do cilindro fixo? Naturalmente, isto se dá no instante em que $N = 0$. Suponhamos que o cilindro comece a cair a partir do repouso ($\dot{\theta} = 0$) da posição vertical ($\theta = 0$). Como não há atrito de deslizamento, a energia mecânica se conserva:

$$T+V=\frac{m}{2}(\dot{r}^2+r^2\dot{\theta}^2)+\frac{ma^2}{4}\dot{\phi}^2+mgr\cos\theta=\frac{3}{4}m(a+b)^2\dot{\theta}^2+mg(a+b)\cos\theta=E\ , \quad (2.4.47)$$

onde foram usados os vínculos. Uma vez que no instante inicial $E=mg(a+b)$, deduzimos

$$\frac{3}{4}(a+b)\dot{\theta}^2=g(1-\cos\theta)\ . \quad (2.4.48)$$

Em decorrência deste último resultado, ficamos com

$$N=\frac{mg}{3}(7\cos\theta-4)\ . \quad (2.4.49)$$

O cilindro móvel abandona o cilindro fixo no ângulo $\theta_0=\cos^{-1}(4/7)$, ou seja, $\theta_0=55^o09'$, resultado independente dos raios dos cilindros.[9] A partir do contato desfeito, os vínculos deixam de existir e o cilindro móvel torna-se um sistema com três graus de liberdade, com as coordenadas r,θ,ϕ independentes entre si. ■

No caso holônomo é possível escolher coordenadas generalizadas em termos das quais as equações de vínculo são identicamente satisfeitas. No entanto, na discussão de certas questões teóricas gerais é freqüentemente vantajoso usar um número excessivo de coordenadas e tratar os vínculos holônomos pela técnica dos multiplicadores de Lagrange. Dados os vínculos holônomos (2.4.30), considere a nova lagrangiana $\mathcal{L}$ definida por

$$\mathcal{L}=L(q,\dot{q},t)+\sum_{l=1}^{p}\lambda_l\, f_l(q,t)\ . \quad (2.4.50)$$

Introduzindo as $(n+p)$ variáveis $\xi_1,\ldots,\xi_{n+p}$ definidas por

$$\xi_1=q_1,\ldots,\xi_n=q_n,\xi_{n+1}=\lambda_1,\ldots,\xi_{n+p}=\lambda_p\ , \quad (2.4.51)$$

e tratando-as como $(n+p)$ coordenadas generalizadas independentes, as equações de movimento e as equações de vínculo podem ser escritas na forma unificada

$$\frac{d}{dt}\Big(\frac{\partial\mathcal{L}}{\partial\dot{\xi}_k}\Big)-\frac{\partial\mathcal{L}}{\partial\xi_k}=0\qquad,\qquad k=1,\ldots,n+p\ . \quad (2.4.52)$$

■ **Exercício 2.4.4.** Prove que as n primeiras equações (2.4.52) são idênticas às equações (2.4.9) com os a_{lk} dados por (2.4.32), ao passo que as p últimas equações (2.4.52) reproduzem as equações de vínculo (2.4.30). ■

As Eqs.(2.4.50) e (2.4.52) mostram que o princípio variacional correto para se formular as equações de movimento de um sistema descrito pela lagrangiana $L(q,\dot{q},t)$ e sujeito aos vínculos holônomos (2.4.30) é

[9] A análise acima não é plenamente realista porque a força de vínculo tangencial tornar-se-á excessivamente pequena para impedir o deslizamento antes de ser atingido o ponto em que $N=0$ (Problema 2.25).

$$\delta \int_{t_1}^{t_2} \left\{ L(q,\dot{q},t) + \sum_{l=1}^{p} \lambda_l \, f_l(q,t) \right\} dt = 0 \ , \tag{2.4.53}$$

onde as coordenadas $q_1, \ldots, q_n$ e os multiplicadores de Lagrange $\lambda_1, \ldots, \lambda_n$ são variados independentemente.

2.5 Propriedades de Simetria e Leis de Conservação

Uma descoberta notável, que adquiriu uma importância fundamental ao longo do século XX, foi a da existência de uma íntima relação entre as leis de conservação e a invariância dos sistemas físicos sob operações de simetria.

Constantes de Movimento

Constantes de movimento são grandezas físicas conservadas, isto é, quantidades associadas a um sistema mecânico que não mudam de valor durante a evolução dinâmica do sistema. Matematicamente, uma *constante de movimento*, *integral primeira* ou simplesmente *integral* de um sistema mecânico é uma função f das coordenadas e velocidades generalizadas, e possivelmente do tempo, que permanece constante durante o movimento do sistema:

$$\frac{df}{dt} = 0 \qquad \text{ou} \qquad f(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t) = \text{constante} \ . \tag{2.5.1}$$

■ **Exemplo 2.5.1.** No caso de um oscilador harmônico simples de freqüência angular ω,

$$f(x,\dot{x},t) = \arctan\left(\frac{\omega\, x}{\dot{x}}\right) - \omega\, t$$

é constante de movimento. A fim de comprovar isto, precisamos calcular a derivada *total* de f relativamente ao tempo, que é dada por

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial \dot{x}}\frac{d\dot{x}}{dt} + \frac{\partial f}{\partial t} = \frac{1}{1+\omega^2 x^2/\dot{x}^2}\,\frac{\omega}{\dot{x}}\,\dot{x} + \frac{1}{1+\omega^2 x^2/\dot{x}^2}\left[-\frac{\omega\, x}{\dot{x}^2}\right]\ddot{x} - \omega \ .$$

Usando a equação de movimento do oscilador, $\ddot{x} = -\omega^2 x$, esta última equação reduz-se a

$$\frac{df}{dt} = \omega\,\frac{1+\omega^2 x^2/\dot{x}^2}{1+\omega^2 x^2/\dot{x}^2} - \omega = \omega - \omega = 0 \ ,$$

e, como antecipado, f é de fato constante de movimento. ■

As constantes de movimento produzem equações diferenciais de primeira ordem que dão importantes informações a respeito da evolução temporal do sistema. Freqüentemente, embora não se conheça a solução exata das equações de movimento, o conhecimento de certas integrais primeiras permite extrair muitas informações fisicamente significativas quanto à natureza do movimento do sistema. É possível, às vezes, obter respostas exatas para certas perguntas graças às constantes de movimento, sem a necessidade da resolução completa do problema dinâmico. Conseqüentemente, investigar condições gerais que assegurem a existência de constantes de movimento é de alta relevância. O

primeiro e mais simples resultado dessa investigação será enunciado após duas definições convenientes.

Definição 2.5.1. Seja $L(q, \dot{q}, t)$ a lagrangiana de um sistema com n graus de liberdade. A quantidade p_k definida por

$$p_k = \frac{\partial L}{\partial \dot{q}_k} \tag{2.5.2}$$

é chamada de *momento conjugado* (ou *momento canônico conjugado*) à coordenada q_k.

■ **Exemplo 2.5.2.** Seja $L = (m/2)(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - V(x, y, z)$. Neste caso, $p_x = m\dot{x}$ é a componente x do momento linear da partícula, com resultados análogos para p_y e p_z. ■

■ **Exemplo 2.5.3.** Seja $L = (m/2)\dot{\mathbf{r}}^2 - e\phi(\mathbf{r}, t) + (e/c)\dot{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r}, t)$ a lagrangiana de uma partícula de carga e num campo eletromagnético externo. Então $p_x = m\dot{x} + (e/c)A_x(\mathbf{r}, t)$, que não coincide com a componente x do momento linear da partícula devido ao termo adicional $(e/c)A_x$. Resultados análogos valem para as componentes p_y e p_z. ■

Definição 2.5.2. Se a lagrangiana de um sistema não contém uma determinada coordenada q_k (embora contenha $\dot{q}_k$), então a referida coordenada é dita *cíclica*[10] ou *ignorável*.

Teorema 2.5.1. O momento conjugado a uma coordenada cíclica é constante de movimento.

Demonstração. Seja q_k uma coordenada cíclica de L. A k-ésima equação de Lagrange

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = 0$$

reduz-se a

$$\frac{dp_k}{dt} = 0$$

em virtude de (2.5.2) e porque $\partial L/\partial q_k = 0$ para uma coordenada cíclica. Portanto, $p_k =$ constante. □

■ **Exemplo 2.5.4.** Em coordenadas esféricas, a lagrangiana de uma partícula sujeita a uma força central escreve-se

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \operatorname{sen}^2\theta\, \dot{\varphi}^2) - V(r) \ . \tag{2.5.3}$$

Claramente φ é coordenada cíclica, donde

[10] Em problemas de grande importância física as coordenadas que não aparecem na lagrangiana costumam ser ângulos, daí chamar-se tais coordenadas genericamente de cíclicas.

$$p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = mr^2 \operatorname{sen}^2\theta\,\dot{\varphi} = \text{constante} \ . \tag{2.5.4}$$

É fácil verificar que p_φ é a componente z do momento angular da partícula. ■

■ **Exemplo 2.5.5.** Um pêndulo oscila suspenso por um ponto que pode deslizar horizontalmente sem atrito. Como vimos no Exemplo 2.3.1, a lagrangiana do sistema é

$$L = \frac{m}{2}(\dot{x}^2 + l^2\dot{\theta}^2 + 2l\dot{x}\dot{\theta}\cos\theta) + mgl\cos\theta \ .$$

Uma vez que x é coordenada cíclica, temos

$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x} + ml\dot{\theta}\cos\theta = \text{constante} \ .$$

Claramente, p_x é a componente x do momento linear da partícula oscilante de massa m. ■

A ausência de uma certa coordenada pode ser interpretada como uma propriedade de simetria da lagrangiana. De fato, se q_k é coordenada cíclica uma alteração no valor de q_k não modifica a lagrangiana. Em outras palavras, a lagrangiana é invariante sob o deslocamento de uma coordenada cíclica. Mas um deslocamento de uma coordenada cíclica de posição corresponde a uma translação, ao passo que o deslocamento de uma coordenada cíclica angular corresponde a uma rotação. No exemplo anterior, a conservação da componente x do momento linear é conseqüência da invariância da lagrangiana frente a translações ao longo da direção x. Correspondentemente, no Exemplo 2.5.4 a conservação da componente z do momento angular decorre da invariância da lagrangiana sob rotações em torno do eixo z. Estas observações sugerem a existência de uma conexão geral entre simetrias da lagrangiana sob translações e rotações e as leis de conservação do momento linear e do momento angular.

Translações e Rotações Infinitesimais

Considere a transformação infinitesimal

$$\mathbf{r}_i \rightarrow \mathbf{r}_i' = \mathbf{r}_i + \delta\mathbf{r}_i \ , \tag{2.5.5a}$$

$$\mathbf{v}_i \rightarrow \mathbf{v}_i' = \mathbf{v}_i + \delta\mathbf{v}_i \ , \tag{2.5.5b}$$

onde $\delta\mathbf{r}_i$ e $\delta\mathbf{v}_i$ são deslocamentos infinitesimais das posições e velocidades de um sistema de n partículas. Seja

$$L(\mathbf{r}_1, \ldots, \mathbf{r}_N, \dot{\mathbf{r}}_1, \ldots, \dot{\mathbf{r}}_N, t) \equiv L(\mathbf{r}, \mathbf{v}, t) \tag{2.5.6}$$

a lagrangiana do sistema. A variação de L sob a transformação (2.5.5) é definida por

$$\delta L = L(\mathbf{r}', \mathbf{v}', t) - L(\mathbf{r}, \mathbf{v}, t) = \sum_{i=1}^{N}\left[\frac{\partial L}{\partial \mathbf{r}_i}\cdot\delta\mathbf{r}_i + \frac{\partial L}{\partial \mathbf{v}_i}\cdot\delta\mathbf{v}_i\right] \ , \tag{2.5.7}$$

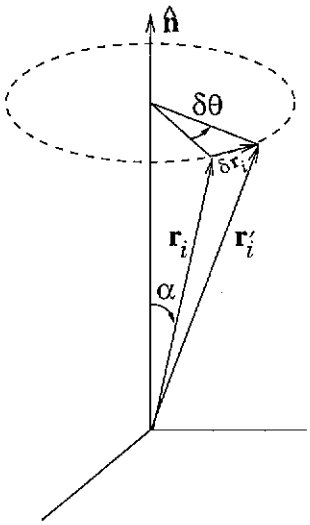


Figura 2.7: Rotação rígida de um sistema de partículas.

onde usamos a notação

$$\frac{\partial L}{\partial \mathbf{r}_i} \equiv \frac{\partial L}{\partial x_i}\hat{\mathbf{i}} + \frac{\partial L}{\partial y_i}\hat{\mathbf{j}} + \frac{\partial L}{\partial z_i}\hat{\mathbf{k}} \ , \tag{2.5.8}$$

com definição análoga para $\partial L/\partial \mathbf{v}_i$.

■ **Exemplo 2.5.6.** Uma translação rígida do sistema de partículas consiste num mesmo deslocamento $\epsilon = \epsilon\hat{\mathbf{n}}$ de todas as partículas do sistema de tal modo que as posições relativas das partículas permaneçam inalteradas, bem como suas velocidades. Neste caso, as Eqs.(2.5.5) são válidas com

$$\delta\mathbf{r}_i = \epsilon\hat{\mathbf{n}} \quad , \quad \delta\mathbf{v}_i = 0 \quad , \quad \text{(translações)} \tag{2.5.9}$$

onde ϵ é um parâmetro infinitesimal. ■

■ **Exemplo 2.5.7.** Uma rotação rígida do sistema de partículas consiste numa rotação de todos os vetores do sistema de um mesmo ângulo $\delta\theta$ em torno de um eixo definido pelo vetor unitário $\hat{\mathbf{n}}$. Inspecionando a Fig. 2.7 conclui-se que $|\delta\mathbf{r}_i| = r_i \operatorname{sen}\alpha\, \delta\theta$, que tem a aparência de módulo de um produto vetorial. De fato, definindo o vetor $\boldsymbol{\delta\theta} = \delta\theta\hat{\mathbf{n}}$ resulta imediatamente que a expressão correta para o vetor $\delta\mathbf{r}_i$ é $\delta\mathbf{r}_i = \boldsymbol{\delta\theta} \times \mathbf{r}_i$. Já que cada velocidade $\delta\mathbf{v}_i$ sofre a mesma rotação, as Eqs.(2.5.5) são válidas com

$$\delta\mathbf{r}_i = \boldsymbol{\delta\theta} \times \mathbf{r}_i \quad , \quad \delta\mathbf{v}_i = \boldsymbol{\delta\theta} \times \mathbf{v}_i \quad , \quad \text{(rotações)} \tag{2.5.10}$$

sendo $\delta\theta$ o parâmetro infinitesimal associado à transformação. ■

Teoremas de Conservação

A fim de formular os resultados numa forma suficientemente geral, suponhamos que o sistema descrito pela lagrangiana (2.5.6) esteja sujeito aos p vínculos holônomos

$$f_s(\mathbf{r}_1, \ldots, \mathbf{r}_N, t) = 0 \quad , \quad s = 1, \ldots, p \ . \tag{2.5.11}$$

Teorema 2.5.2. Suponha que um sistema mecânico seja descrito pela lagrangiana $L = T - V$, onde V é um potencial independente das velocidades. Se a lagrangiana e os vínculos (2.5.11) são invariantes sob uma translação rígida arbitrária, então o momento linear total do sistema é conservado.

Demonstração. Conforme a discussão ao final da Seção 2.4, a lagrangiana

$$\mathcal{L} = L(\mathbf{r}, \mathbf{v}, t) + \sum_{s=1}^{p} \lambda_s f_s(\mathbf{r}, t) \tag{2.5.12}$$

fornece de um só golpe as equações de movimento e as equações de vínculo. Uma vez que L e as equações de vínculo são invariantes sob translações arbitrárias, $\mathcal{L}$ também o é. Em particular, $\mathcal{L}$ é invariante sob qualquer translação infinitesimal, donde

$$0 = \delta\mathcal{L} = \sum_{i=1}^{N} \frac{\partial \mathcal{L}}{\partial \mathbf{r}_i} \cdot (\epsilon \hat{\mathbf{n}}) = \epsilon \hat{\mathbf{n}} \cdot \sum_{i=1}^{N} \frac{\partial \mathcal{L}}{\partial \mathbf{r}_i} \ , \tag{2.5.13}$$

onde usamos (2.5.7) e (2.5.9). Como ϵ é arbitrário, deduzimos

$$\hat{\mathbf{n}} \cdot \sum_{i=1}^{N} \frac{\partial \mathcal{L}}{\partial \mathbf{r}_i} = 0 \ . \tag{2.5.14}$$

Mas, notando que

$$\frac{\partial \mathcal{L}}{\partial \mathbf{v}_i} \equiv \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_i} = \frac{\partial L}{\partial \dot{\mathbf{r}}_i} = m_i \dot{\mathbf{r}}_i = \mathbf{p}_i \ , \tag{2.5.15}$$

onde $\mathbf{p}_i$ é o momento linear da i-ésima partícula, as equações de Lagrange podem ser escritas na forma

$$\frac{\partial \mathcal{L}}{\partial \mathbf{r}_i} = \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_i}\right) = \frac{d\mathbf{p}_i}{dt} \ . \tag{2.5.16}$$

Assim, a Eq.(2.5.14) torna-se

$$\frac{d}{dt}\Big[\hat{\mathbf{n}} \cdot \sum_{i=1}^{N} \mathbf{p}_i\Big] \equiv \frac{d}{dt}\Big(\hat{\mathbf{n}} \cdot \mathbf{P}\Big) = 0 \ , \tag{2.5.17}$$

provando que a componente do momento linear total $\mathbf{P}$ ao longo da direção $\hat{\mathbf{n}}$ é conservada. Escolhendo sucessivamente $\hat{\mathbf{n}} = \hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$, conclui-se que P_x, P_y e P_z são constantes de movimento, isto é, o vetor momento linear total é conservado. □

Teorema 2.5.3. Seja V um potencial independente das velocidades. Se a lagrangiana $L = T - V$ e os vínculos (2.5.11) são invariantes sob uma rotação rígida arbitrária, então o momento angular total do sistema é conservado.

Demonstração. A invariância de $\mathcal{L}$ sob qualquer rotação infinitesimal implica

$$0 = \delta\mathcal{L} = \sum_{i=1}^{N}\left[\frac{\partial\mathcal{L}}{\partial\mathbf{r}_i}\cdot(\boldsymbol{\delta\theta}\times\mathbf{r}_i) + \frac{\partial\mathcal{L}}{\partial\mathbf{v}_i}\cdot(\boldsymbol{\delta\theta}\times\mathbf{v}_i)\right] , \tag{2.5.18}$$

onde usamos (2.5.7) e (2.5.10). O produto misto é invariante sob uma permutação cíclica dos vetores envolvidos, o que permite escrever

$$\delta\theta\,\hat{\mathbf{n}}\cdot\sum_{i=1}^{N}\left[\mathbf{r}_i\times\frac{\partial\mathcal{L}}{\partial\mathbf{r}_i} + \mathbf{v}_i\times\frac{\partial\mathcal{L}}{\partial\mathbf{v}_i}\right] = 0 . \tag{2.5.19}$$

Tendo em conta (2.5.15), as equações de Lagrange (2.5.16) e a arbitrariedade de $\delta\theta$, somos conduzidos a

$$\hat{\mathbf{n}}\cdot\sum_{i=1}^{N}(\mathbf{r}_i\times\dot{\mathbf{p}}_i + \dot{\mathbf{r}}_i\times\mathbf{p}_i) = \frac{d}{dt}\left[\hat{\mathbf{n}}\cdot\sum_{i=1}^{N}\mathbf{r}_i\times\mathbf{p}_i\right] \equiv \frac{d}{dt}(\hat{\mathbf{n}}.\mathbf{L}) = 0 . \tag{2.5.20}$$

Portanto, a componente do momento angular total $\mathbf{L}$ ao longo da direção $\hat{\mathbf{n}}$ é constante de movimento. A arbitrariedade de $\hat{\mathbf{n}}$ implica a conservação do vetor momento angular total. □

Os dois últimos resultados aplicam-se a sistemas de partículas isolados e, também, a sistemas de corpos rígidos isolados porque, neste caso, os vínculos de rigidez – distâncias fixas entre todos os pares de pontos – são obviamente preservados por rotações ou translações. É importante sublinhar, ainda, que um dado sistema pode ser dotado de simetria parcial sob rotações ou translações, o que acarreta a conservação de algumas das componentes de $\mathbf{P}$ ou $\mathbf{L}$, mas não de todas. A energia cinética é invariante sob translações ou rotações arbitrárias, de modo que a invariância da lagrangiana é determinada pelas simetrias do potencial. Considere, por exemplo, uma partícula movendo-se no potencial gravitacional de um plano homogêneo infinito. O potencial é claramente invariante sob translações paralelas ao plano e sob rotações em torno de um eixo perpendicular ao plano (que tomaremos como eixo z). Conseqüentemente, a lagrangiana é invariante sob translações ao longo das direções x e y, e sob rotações em torno do eixo z. Assim, P_x, P_y e L_z são quantidades conservadas.

Incidentalmente, as Eqs.(2.5.17) e (2.5.20) mostram que, para um sistema isolado, portanto invariante sob translações e rotações, a soma das forças internas e dos torques internos é zero. É notável que a simetria sob translações e rotações garante a conservação do momento linear total e do momento angular total, sem a necessidade de se invocar a terceira lei de Newton nem sequer em sua forma fraca.

2.6 Conservação da Energia

Naturalmente, é de se esperar que a dinâmica lagrangiana contenha o teorema da conservação da energia para sistemas cujas forças derivam de um potencial dependente apenas das posições das partículas. Na verdade, é possível demonstrar um teorema de

conservação bastante geral que se reduz à lei da conservação da energia num caso especial. Embora tenha sido vantajoso o emprego de coordenadas cartesianas para lidar com translações e rotações, aqui será mais conveniente considerar o sistema descrito por coordenadas generalizadas $q_1, \ldots, q_n$ e pela lagrangiana $L(q, \dot{q}, t)$. Suporemos que o sistema é holônomo, de modo que as coordenadas generalizadas são efetivamente independentes entre si. A generalização para o caso não-holônomo é deixada como um exercício (Problema 2.24).

Teorema 2.6.1. Sejam $q_1, \ldots, q_n$ coordenadas generalizadas de um sistema com lagrangiana L. Se L não depende *explicitamente* do tempo, a quantidade h definida por[11]

$$h = \sum_{k=1}^{n} \dot{q}_k \frac{\partial L}{\partial \dot{q}_k} - L \tag{2.6.1}$$

é uma constante de movimento, freqüentemente chamada de *integral de Jacobi.*

Demonstração. A derivada total de h em relação ao tempo é

$$\frac{dh}{dt} = \sum_{k=1}^{n} \left\{ \ddot{q}_k \frac{\partial L}{\partial \dot{q}_k} + \dot{q}_k \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) \right\} - \left\{ \sum_{k=1}^{n} \left(\frac{\partial L}{\partial q_k}\dot{q}_k + \frac{\partial L}{\partial \dot{q}_k}\ddot{q}_k \right) + \frac{\partial L}{\partial t} \right\} ,$$

donde, usando as equações de Lagrange,

$$\frac{dh}{dt} = -\frac{\partial L}{\partial t} . \tag{2.6.2}$$

Se L não depende explicitamente do tempo, $\partial L/\partial t = 0$ e a Eq.(2.6.2) estabelece que h é constante de movimento. □

Sob circunstâncias bem definidas, que passamos a examinar, h é a energia total do sistema. A energia cinética, expressa em termos das coordenadas e velocidades generalizadas, escreve-se

$$T = \frac{1}{2}\sum_i m_i v_i^2 = \frac{1}{2}\sum_i m_i \left(\sum_k \frac{\partial \mathbf{r}_i}{\partial q_k}\dot{q}_k + \frac{\partial \mathbf{r}_i}{\partial t} \right) \cdot \left(\sum_l \frac{\partial \mathbf{r}_i}{\partial q_l}\dot{q}_l + \frac{\partial \mathbf{r}_i}{\partial t} \right) , \tag{2.6.3}$$

tendo sido usada a Eq.(1.4.4). Esta última equação pode ser reescrita na forma

$$T = M_0 + \sum_k M_k \dot{q}_k + \frac{1}{2}\sum_{k,l} M_{kl}\dot{q}_k\dot{q}_l , \tag{2.6.4}$$

onde M_0, M_k e M_{kl} são funções dos qs e do tempo dadas por

$$M_0 = \frac{1}{2}\sum_i m_i \frac{\partial \mathbf{r}_i}{\partial t} \cdot \frac{\partial \mathbf{r}_i}{\partial t} , \tag{2.6.5a}$$

$$M_k = \sum_i m_i \frac{\partial \mathbf{r}_i}{\partial t} \cdot \frac{\partial \mathbf{r}_i}{\partial q_k} , \tag{2.6.5b}$$

[11] Às vezes h é chamada de "função energia" ("energy function", em inglês).

$$M_{kl} = \sum_i m_i \frac{\partial \mathbf{r}_i}{\partial q_k} \cdot \frac{\partial \mathbf{r}_i}{\partial q_l} \ . \tag{2.6.5c}$$

■ **Exercício 2.6.1.** Deduza as Eqs.(2.6.5). ■

Se $L = T - V$, com V independente das velocidades, tem-se

$$h = \sum_k \frac{\partial T}{\partial \dot{q}_k} \dot{q}_k - (T - V) \ . \tag{2.6.6}$$

Se, além disso, as Eqs.(1.5.2) que definem as coordenadas generalizadas não envolvem explicitamente o tempo, $\partial \mathbf{r}_i / \partial t = 0$, de modo que $M_0 = M_k = 0$ e (2.6.4) nos diz que T é uma função homogênea do segundo grau das velocidades generalizadas. Portanto, pelo teorema de Euler das funções homogêneas (Apêndice B),

$$h = 2T - (T - V) = T + V = E \ , \tag{2.6.7}$$

e h é a energia total do sistema. Nas presentes circunstâncias, se V não é função explícita do tempo L tampouco o é, de modo que a energia total — no caso, igual a h — será conservada. Vale notar que, mesmo na ausência de potenciais com dependência temporal explícita, a existência de vínculos dependentes do tempo (como no Exemplo 1.5.2) introduz uma dependência temporal explícita nas equações (1.4.2), fazendo com que h não coincida com a energia total. Cumpre destacar, ainda, que as condições que asseguram a conservação de h são completamente independentes daquelas que garantem a igualdade entre h e a energia total. Assim, h pode ser conservada sem ser a energia total, ou ser a energia total sem ser conservada. Por outro lado, ao contrário da lagrangiana, cujo valor independe da escolha das coordenadas generalizadas, o valor e a forma funcional de h dependem do particular conjunto de coordenadas generalizadas que se tenha escolhido.

De modo geral, quando há vínculos que dependem explicitamente do tempo a energia total não se conserva, e a razão é simples. Como vimos na Seção 1.3, se os vínculos mudam com o tempo as forças de vínculo realizam trabalho por ocasião de deslocamentos *reais*, embora o trabalho realizado durante deslocamentos virtuais seja nulo. Ora, a variação da energia total num intervalo de tempo dt é igual ao trabalho realizado pelas forças de vínculo durante os deslocamentos reais $d\mathbf{r}_i = \mathbf{v}_i dt$, que em geral não se anula, logo a energia total não pode permanecer constante.

Os três últimos teoremas lançam uma nova luz sobre a origem das constantes de movimento fundamentais da mecânica, ao associarem as leis de conservação a propriedades geométricas do espaço-tempo. A homogeneidade e isotropia do espaço refletem-se na invariância da lagrangiana e dos vínculos sob translações e rotações, dando lugar à conservação do momento linear e do momento angular. A homogeneidade temporal reflete-se na imutabilidade da lagrangiana e dos vínculos frente a uma mudança na origem do tempo (lagrangiana e vínculos sem dependência temporal explícita), dando lugar à conservação da energia.

Quando há forças dissipativas presentes que podem ser descritas por uma função de dissipação $\mathcal{F}$, o Teorema 2.6.1 não é mais verdadeiro e a Eq.(2.6.2) é substituída por

$$\frac{dh}{dt} = -\sum_k \dot{q}_k \frac{\partial \mathcal{F}}{\partial \dot{q}_k} - \frac{\partial L}{\partial t} \ . \tag{2.6.8}$$

■ **Exercício 2.6.2.** Deduza a Eq.(2.6.8). ■

Note que $\mathcal{F}$, no que concerne à sua dependência funcional das velocidades generalizadas, tem a mesma estrutura que a energia cinética. Assim, se as Eqs.(1.4.2) que definem as coordenadas generalizadas não dependem explicitamente do tempo, $h = E$ e $\mathcal{F}$ é uma função homogênea do segundo grau das velocidades. Neste caso (Lemos 1991), o teorema de Euler das funções homogêneas permite escrever (2.6.8) na forma

$$\frac{dE}{dt} = -2\mathcal{F} - \frac{\partial L}{\partial t} \ . \tag{2.6.9}$$

Se, em acréscimo às circunstâncias já especificadas, L não depende explicitamente do tempo, vale

$$\frac{dE}{dt} = -2\mathcal{F} \ , \tag{2.6.10}$$

já obtida por outros métodos, porém num contexto menos geral, na Seção 1.6.

2.7 Teorema de Noether

Na dinâmica lagrangiana a conexão geral entre propriedades de simetria (invariância) e quantidades conservadas é estabelecida por um importante teorema devido a Emmy Noether. Esse teorema (Noether 1918) contém como casos particulares todos os resultados discutidos nas Seções 2.5 e 2.6.

Sejam X e Ψ_i funções conhecidas de $(n+1)$ variáveis reais e seja ϵ um parâmetro infinitesimal arbitrário. Considere a transformação infinitesimal

$$t \longrightarrow t' = t + \epsilon X(q(t), t) \ , \tag{2.7.1a}$$

$$q_i(t) \longrightarrow q_i'(t') = q_i(t) + \epsilon \Psi_i(q(t), t) \ . \tag{2.7.1b}$$

Dizemos que a integral de ação permanece invariante sob esta transformação se

$$\Delta S = \int_{t_1'}^{t_2'} L\Big(q'(t'), \frac{dq'(t')}{dt'}, t'\Big) dt' - \int_{t_1}^{t_2} L\Big(q(t), \frac{dq(t)}{dt}, t\Big) dt = 0 \ . \tag{2.7.2}$$

Teorema de Noether. Dado um sistema mecânico com n graus de liberdade, se a ação é invariante sob a transformação (2.7.1) então a quantidade

$$C = \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i}(\dot{q}_i X - \Psi_i) - LX \tag{2.7.3}$$

é constante de movimento, onde $L(q,\dot{q},t)$ é a lagrangiana do sistema.

Demonstração. Nos cálculos que se seguem reteremos somente termos de primeira ordem em ϵ. Notemos inicialmente que

$$\frac{dt'}{dt} = 1 + \epsilon\dot{X} \quad , \quad \frac{dt}{dt'} = (1+\epsilon\dot{X})^{-1} = 1 - \epsilon\dot{X} \ , \tag{2.7.4}$$

tendo sido usada a expansão binomial $(1+x)^\lambda = 1 + \lambda x + \cdots$. Temos, ainda,

$$\frac{dq_i'(t')}{dt'} = \frac{dt}{dt'}\frac{dq_i'(t')}{dt} = (1-\epsilon\dot{X})(\dot{q}_i + \epsilon\dot{\Psi}_i) = \dot{q}_i + \epsilon\xi_i \ , \tag{2.7.5}$$

onde

$$\xi_i = \dot{\Psi}_i - \dot{q}_i\dot{X} \ . \tag{2.7.6}$$

Assim, a invariância da ação, isto é,

$$\begin{aligned}\Delta S &= \int_{t_1}^{t_2} L(q+\epsilon\Psi, \dot{q}+\epsilon\xi, t+\epsilon X)(1+\epsilon\dot{X})dt - \int_{t_1}^{t_2} L(q,\dot{q},t)dt \\ &= \int_{t_1}^{t_2}\left\{L(q,\dot{q},t) + \sum_{i=1}^{n}\Big[\frac{\partial L}{\partial q_i}\epsilon\Psi_i + \frac{\partial L}{\partial \dot{q}_i}\epsilon\xi_i\Big] + \frac{\partial L}{\partial t}\epsilon X\right\}(1+\epsilon\dot{X})dt - \int_{t_1}^{t_2} L(q,\dot{q},t)dt \\ &= \epsilon\int_{t_1}^{t_2}\left\{\sum_{i=1}^{n}\Big[\frac{\partial L}{\partial q_i}\Psi_i + \frac{\partial L}{\partial \dot{q}_i}\xi_i\Big] + \frac{\partial L}{\partial t}X + L\dot{X}\right\}dt = 0\end{aligned} \tag{2.7.7}$$

conduz imediatamente à condição de Noether

$$\sum_{i=1}^{n}\left[\Psi_i\frac{\partial L}{\partial q_i} + (\dot{\Psi}_i - \dot{q}_i\dot{X})\frac{\partial L}{\partial \dot{q}_i}\right] + L\dot{X} + \frac{\partial L}{\partial t}X = 0 \ , \tag{2.7.8}$$

já que ϵ e o intervalo de integração são arbitrários. *Esta é a condição necessária e suficiente para que a ação seja invariante frente à transformação* (2.7.1). Com o emprego das equações de Lagrange e das Eqs.(2.6.1) e (2.6.2), a condição de Noether (2.7.8) pode ser reescrita na forma

$$\sum_{i=1}^{n}\left[\Psi_i\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_i}\Big) + \dot{\Psi}_i\frac{\partial L}{\partial \dot{q}_i}\right] - h\dot{X} - \frac{dh}{dt}X = \frac{d}{dt}\Big\{\sum_{i=1}^{n}\frac{\partial L}{\partial \dot{q}_i}\Psi_i - hX\Big\} = 0 \tag{2.7.9}$$

ou, equivalentemente,

$$\frac{d}{dt}\left\{\sum_{i=1}^{n}\frac{\partial L}{\partial \dot{q}_i}(\dot{q}_iX - \Psi_i) - LX\right\} = 0 \ , \tag{2.7.10}$$

o que completa a demonstração do teorema. □

Generalização do Teorema de Noether

Segundo o Teorema 2.3.1, lagrangianas que só diferem pela derivada total em relação ao tempo de uma função arbitrária das coordenadas generalizadas e do tempo conduzem exatamente às mesmas equações de movimento. O teorema de Noether admite uma generalização importante que reflete esta propriedade das lagrangianas equivalentes. Suponha que a integral de ação seja quase-invariante sob a transformação infinitesimal (2.7.1), isto é, imagine que exista uma função $G(q,t)$ tal que

$$\Delta S = \int_{t_1'}^{t_2'} L\Big(q'(t'), \frac{dq'(t')}{dt'}, t'\Big)dt' - \int_{t_1}^{t_2} \left\{ L\Big(q(t), \frac{dq(t)}{dt}, t\Big) + \epsilon \frac{d}{dt} G(q(t),t) \right\} dt = 0 \ . \tag{2.7.11}$$

Então a condição de Noether assume a forma

$$\sum_{i=1}^{n} \left[\Psi_i \frac{\partial L}{\partial q_i} + (\dot{\Psi}_i - \dot{q}_i \dot{X}) \frac{\partial L}{\partial \dot{q}_i} \right] + L\dot{X} + \frac{\partial L}{\partial t} X = \dot{G} \tag{2.7.12}$$

e a quantidade conservada associada à simetria é

$$\bar{C} = \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i} (\dot{q}_i X - \Psi_i) - LX + G \ . \tag{2.7.13}$$

Os detalhes da demonstração ficam por conta do leitor.

Algumas Aplicações Simples

As leis de conservação fundamentais são casos particulares do teorema de Noether, o que revela o seu poderio e sua generalidade.

■ **Exemplo 2.7.1.** Demonstrar o Teorema 2.6.1 por meio do teorema de Noether.

Solução: Considere um deslocamento puramente temporal, ou seja, tome a transformação (2.7.1) com $\Psi_i = 0$ e $X = 1$. De acordo com a condição de Noether (2.7.8), a ação será invariante sob esta transformação se e somente se $\partial L/\partial t = 0$. Se este for o caso, por (2.7.3) concluímos que

$$h = \sum_{i=1}^{n} \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L = \text{constante} \ , \tag{2.7.14}$$

o que demonstra o Teorema (2.6.1). ■

■ **Exemplo 2.7.2.** Demonstrar o Teorema 2.5.3 recorrendo ao teorema de Noether.

Solução: Defina $q_1 = x_1, q_2 = y_1, q_3 = z_1, \ldots, q_{3N} = z_N$ e tome $\delta\theta = \epsilon\hat{\mathbf{n}}$, com $\epsilon = \delta\theta$. As Eqs.(2.5.10) são da forma (2.7.1) com $X = 0$, $\Psi_1 = (\hat{\mathbf{n}} \times \mathbf{r}_1)_x, \Psi_2 = (\hat{\mathbf{n}} \times \mathbf{r}_1)_y, \Psi_3 = (\hat{\mathbf{n}} \times \mathbf{r}_1)_z, \ldots, \Psi_{3N} = (\hat{\mathbf{n}} \times \mathbf{r}_N)_z$. Por outro lado, a invariância de S deve-se à invariância de $\mathcal{L}$, pois não há mudança da variável temporal. Assim, se $\mathcal{L}$ é invariante o teorema de Noether garante que

$$-\sum_{k=1}^{3N} \frac{\partial \mathcal{L}}{\partial \dot{q}_k} \Psi_k \equiv -\sum_{i=1}^{N} \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_i} \cdot (\hat{\mathbf{n}} \times \mathbf{r}_i) = \text{constante} \ ,$$

ou seja,

$$\hat{\mathbf{n}} \cdot \sum_{i=1}^{N} \mathbf{r}_i \times \mathbf{p}_i = \text{constante} \ ,$$

que reproduz a Eq.(2.5.20). ■

■ **Exercício 2.7.1.** Prove o teorema (2.5.2) utilizando o teorema de Noether. ■

Consideremos, por fim, um exemplo original de aplicação do teorema de Noether, que envolve uma transformação conjunta de espaço e tempo.

■ **Exemplo 2.7.3.** (Lemos 1993) A equação de movimento de um oscilador harmônico amortecido pode ser obtida da lagrangiana (Problema 1.14)

$$L = e^{\lambda t}\left(\frac{m}{2}\dot{x}^2 - \frac{m\omega^2}{2}x^2\right) \ . \tag{2.7.15}$$

Mostrar que a ação é invariante sob a transformação *finita* (α é uma constante arbitrária)

$$t' = t + \alpha \quad , \quad x'(t') = e^{-\lambda\alpha/2}\,x(t) \ , \tag{2.7.16}$$

e deduzir que

$$C = e^{\lambda t}\left(\frac{m}{2}\dot{x}^2 + \frac{m\omega^2}{2}x^2 + \frac{m\lambda}{2}x\dot{x}\right) \tag{2.7.17}$$

é constante de movimento.

Solução. Note que (2.7.16) consiste numa translação temporal acompanhada de uma dilatação espacial. Temos

$$S' = \int_{t_1'}^{t_2'} e^{\lambda t'}\left\{\frac{m}{2}\left(\frac{dx'}{dt'}\right)^2 - \frac{m\omega^2}{2}x'^2\right\}dt' = \int_{t_1+\alpha}^{t_2+\alpha} e^{\lambda t}e^{\lambda\alpha}e^{-\lambda\alpha}\left(\frac{m}{2}\dot{x}^2 - \frac{m\omega^2}{2}x^2\right)dt' =$$

$$= \int_{t_1}^{t_2} e^{\lambda t}\left(\frac{m}{2}\dot{x}^2 - \frac{m\omega^2}{2}x^2\right)dt = S \ , \tag{2.7.18}$$

tendo sido feita a mudança de variável de integração $t' = t + \alpha$ na segunda integral em (2.7.18). Portanto, a ação é invariante e o teorema de Noether se aplica. A versão infinitesimal das Eqs.(2.7.16) é obtida fazendo $\alpha = \epsilon$. Expandindo a exponencial até primeira ordem em ϵ, resulta

$$t' = t + \epsilon \quad , \quad x'(t') = x(t) - \frac{\epsilon\lambda}{2}x(t) \ , \tag{2.7.19}$$

donde, por comparação com (2.7.1),

$$X = 1 \quad , \quad \Psi = -\frac{\lambda}{2}x \ . \tag{2.7.20}$$

Substituindo (2.7.20) em (2.7.3) com $q_1 = x$, $\Psi_1 = -\lambda x/2$ e a soma reduzida a um único termo, resulta imediatamente a constante de movimento (2.7.17). ■

■ **Exercício 2.7.2.** (i) Usando (2.7.15) e (2.7.20), demonstre a invariância da ação verificando que a condição de Noether (2.7.8) é satisfeita. (ii) Utilizando a equação de movimento do oscilador harmônico amortecido, prove diretamente que a quantidade C definida por (2.7.17) é constante de movimento. ■

Princípios Variacionais: Física e Metafísica

Como vimos neste capítulo, a formulação de uma lei física na forma de um princípio variacional oferece importantes vantagens. Em primeiro lugar, torna manifesta a invariância da referida lei sob transformações arbitrárias de coordenadas porque os princípios variacionais independem das particulares coordenadas em que estejam expressos. Em segundo lugar, permite estabelecer uma conexão de grande generalidade entre simetrias e leis de conservação. Quanto ao aspecto estético, os princípios variacionais exprimem as leis físicas numa forma incomparavelmente concisa e elegante. Além disso, os princípios variacionais servem de base para cálculos aproximados muito eficientes em problemas que não se consegue resolver exatamente (Problema 2.4; Feynman, Leighton & Sands 1963, Vol. II, Cap. 19).

Por outro lado, ao fazerem referência ao movimento em sua totalidade, os princípios variacionais parecem contrabandear para a Física a noção de propósito ou finalidade, dotando os sistemas mecânicos de características antropomórficas. Será que uma partícula de algum modo "fareja" todas as trajetórias possíveis antes de se decidir por aquela que dá a menor ação? Por suscitar questões dessa natureza, os princípios variacionais, desde sua introdução na mecânica, vêm exercendo um grande fascínio sobre cientistas e filósofos. Ao longo dos séculos, filósofos e religiosos de variados matizes, às vezes apelando para teorias científicas, têm se dedicado a vãs tentativas de demonstrar a existência de Deus. O surpreendente é que até mesmo um físico da estatura de Max Planck, sucumbindo ao misticismo, tenha chegado a advogar uma interpretação teleológica para o princípio da mínima ação e a invocá-lo como argumento para a existência do Ser Supremo. O leitor inclinado a especulações metafísicas é remetido a Yourgrau & Mandelstam (1968) para uma estimulante discussão crítica dessas idéias e das implicações filosóficas dos princípios variacionais.

PROBLEMAS

2.1. Determine as geodésicas de um cone com ângulo de abertura 2α. Descreva a forma da curva geodésica no caso em que seu ponto inicial é o vértice do cone, tomado como origem do sistema de coordenadas.

2.2. Uma partícula move-se num plano vertical sob um campo gravitacional constante. Mostre que existe uma única solução $\mathbf{r}_f(t)$ da equação de movimento com as condições de contorno $\mathbf{r}(0) = 0$, $\mathbf{r}(T) = \mathbf{r}_0$. Seja agora $\mathbf{r}(t) = \mathbf{r}_f(t) + \boldsymbol{\eta}(t)$, com $\boldsymbol{\eta}(T) = \boldsymbol{\eta}(0) = 0$, uma outra trajetória qualquer obedecendo às mesmas condições de contorno. Fazendo uma integração por partes e usando a equação de movimento satisfeita por $\mathbf{r}_f(t)$, mostre que a ação associada à trajetória $\mathbf{r}(t)$ é dada por $S[\mathbf{r}] = S[\mathbf{r}_f] + (m/2)\int_0^T |\dot{\boldsymbol{\eta}}|^2 dt$ e daí conclua que a ação é mínima para a trajetória física.

2.3. Para um oscilador harmônico de massa m e freqüência angular ω, se $x_f(t)$ é solução da equação de movimento e $x(t) = x_f(t) + \eta(t)$ com $\eta(0) = \eta(T) = 0$, procedendo como no problema anterior mostre que $S[x] = S[x_f] + (m/2)\int_0^T (\dot{\eta}^2 - \omega^2\eta^2)dt$. Expandindo η na série de Fourier $\eta(t) = \sum_{n=1}^{\infty} C_n \operatorname{sen}(n\pi t/T)$ (por que isto é possível?), mostre que $S[x] = S[x_f] + (mT/4)\sum_{n=1}^{\infty}(\frac{n^2\pi^2}{T^2} - \omega^2)\,C_n^2$ e conclua que a ação é mínima para a trajetória física se $T < \pi/\omega$.

2.4. Considere o problema de determinar o mínimo de $J[y] = \int_0^1 y'^2 dx$ com $y(0) = 0$ e $y(1) = 1$. Tente uma solução aproximada da forma $y_\alpha(x) = x^3 + \alpha x(x-1)$ e determine o valor α_0 do parâmetro α que fornece a melhor aproximação. Prove que a solução exata é $y(x) = x$. Faça um gráfico comparativo de $y_{\alpha_0}(x)$ e $y(x)$. Compare o valor aproximado do mínimo, isto é, $J[y_{\alpha_0}]$, com o valor exato $J[y]$ e calcule o erro percentual cometido.

2.5. No cálculo das variações, o problema de achar o extremo do funcional

$$J[y] = \int_{x_1}^{x_2} F(y_1(x), \ldots, y_n(x), y_1'(x), \ldots y_n'(x), x)\, dx$$

com as condições subsidiárias

$$\int_{x_1}^{x_2} f_i(y_1(x), \ldots, y_n(x), y_1'(x), \ldots y_n'(x), x)\, dx = \ell_i \quad , \qquad i = 1, \ldots, m \ ,$$

onde os ℓ_i são constantes, é chamado de *problema isoperimétrico*. Pode-se provar (Elsgoltz 1969) que o extremo de $J[y]$ é determinado por meio das equações de Euler para o funcional

$$\bar{J}[y] = \int_{x_1}^{x_2} \Big(F + \sum_{i=1}^{m} \lambda_i f_i\Big)\, dx \ ,$$

onde os multiplicadores de Lagrange λ_i são constantes. Use este resultado para determinar a forma de um cabo de comprimento ℓ, flexível, inextensível e homogêneo, que pende de dois postes verticais. Levando em conta que a posição de equilíbrio do cabo é aquela em que sua energia potencial gravitacional é mínima, mostre que o cabo toma a forma da curva que minimiza $E[y] = \int_a^b y\sqrt{1+y'^2}\, dx$ com a condição $\int_a^b \sqrt{1+y'^2}\, dx = \ell$, e mostre que a solução é uma catenária.

2.6. Mostre que as equações de Lagrange podem ser escritas na forma explícita

$$\sum_{j=1}^{n} \frac{\partial^2 L}{\partial \dot{q}_j \partial \dot{q}_i} \ddot{q}_j + \sum_{j=1}^{n} \frac{\partial^2 L}{\partial q_j \partial \dot{q}_i} \dot{q}_j + \frac{\partial^2 L}{\partial t \partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \ , \quad i = 1, \ldots, n \ .$$

Prove que duas lagrangianas $\bar{L}(q, \dot{q}, t)$ e $L(q, \dot{q}, t)$ geram equações de movimento *idênticas* se e somente se elas diferem pela derivada total em relação ao tempo de uma função arbitrária das coordenadas generalizadas e do tempo. Você deverá necessitar do resultado contido no Apêndice D.

2.7. Mostre que a lagrangiana

$$\bar{L} = \frac{1}{12}\dot{x}^4 + \frac{\omega^2}{2} x^2 \dot{x}^2 - \frac{\omega^4}{4} x^4$$

gera uma equação de movimento equivalente à de um oscilador harmônico de massa $m = 1$, mas não difere meramente por uma derivada total da lagrangiana usual $L = \dot{x}^2/2 - \omega^2 x^2/2$.

2.8. Considere a máquina de Atwood oscilante do Problema 1.5 no caso especial $M = 3m$. Prove diretamente que a quantidade

$$I(r, \dot{r}, \theta, \dot{\theta}) = r^2 \dot{\theta} \left[\dot{r} \cos \frac{\theta}{2} - \frac{r\dot{\theta}}{2} \operatorname{sen} \frac{\theta}{2} \right] + g r^2 \operatorname{sen} \frac{\theta}{2} \cos^2 \frac{\theta}{2}$$

é constante de movimento (Tufillaro 1986).

2.9. Considere um oscilador harmônico com freqüência dependente do tempo, cuja lagrangiana é $L = m\dot{q}^2/2 - m\omega^2(t) q^2/2$. (i) Se a função $\rho(t)$ satisfaz a equação diferencial

$$\ddot{\rho} - \frac{1}{\rho^3} + \omega^2(t)\, \rho = 0 \ ,$$

prove que

$$I(q, \dot{q}, t) = (\rho \dot{q} - \dot{\rho} q)^2 + \frac{q^2}{\rho^2}$$

é constante de movimento (Lewis Jr. & Riesenfeld 1969). (ii) Um exemplo de interesse em cosmologia é $\omega^2(t) = \omega_0^2 + 2/(9t^2)$, com ω_0 constante. Mostre que, neste caso, uma solução para $\rho(t)$ é

$$\rho(t) = \left(\frac{\pi t}{2}\right)^{1/2} \left[J_{1/6}^2(\omega_0 t) + Y_{1/6}^2(\omega_0 t) \right]^{1/2} \ ,$$

onde J_ν e Y_ν são funções de Bessel de primeira e segunda espécies, respectivamente (Lemos & Natividade 1987).

2.10. Uma partícula move-se no potencial gravitacional produzido pelas seguintes distribuições de massa homogêneas: (i) esfera de raio R; (ii) paralelepípedo infinitamente longo com seção transversal quadrada; (iii) haste de comprimento ℓ; (iv) disco de raio R; (v) cilindro infinitamente longo com seção transversal circular; (vi) plano infinito; (vii)

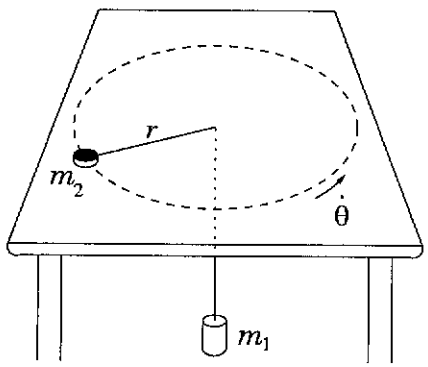


Figura 2.8: Problema 2.12.

plano semi-infinito; (viii) fio enrolado na forma de uma hélice infinita de raio R e passo p. Quais são as componentes de $\mathbf{P}$ e $\mathbf{L}$ que se conservam em cada caso?

2.11. O campo eletromagnético é invariante sob uma transformação de calibre dos potenciais definida por

$$\mathbf{A} \rightarrow \mathbf{A}' = \mathbf{A} + \nabla\Lambda \quad , \quad \phi \rightarrow \phi' = \phi - \frac{1}{c}\frac{\partial\Lambda}{\partial t} \ ,$$

onde $\Lambda(\mathbf{r}, t)$ é uma função diferenciável arbitrária. De que maneira a lagrangiana (1.6.11) é afetada por esta transformação? As equações de movimento são alteradas?

2.12. O sistema representado na figura é tal que a massa m_2 se move sem atrito sobre a mesa horizontal e a massa m_1 só pode movimentar-se verticalmente. Use o método dos multiplicadores de Lagrange para provar que a tensão no fio inextensível que une as partículas é dada por

$$T = \frac{m_1 m_2}{m_1 + m_2}\left[g + \frac{p_\theta^2}{m_2^2 r^3}\right] \ ,$$

onde p_θ é o valor constante de $m_2 r^2\dot{\theta}$. No caso em que $p_\theta = 0$, determine a aceleração de queda da massa m_1.

2.13. Na chamada *mecânica generalizada* as lagrangianas contêm derivadas de ordem superior à primeira das coordenadas generalizadas. Dada uma lagrangiana da forma $L(q, \dot{q}, \ddot{q}, t)$, mostre que o princípio de Hamilton, com variação nula de q_i e $\dot{q}_i$ nos extremos temporais, dá lugar às equações de Lagrange

$$\frac{d^2}{dt^2}\left(\frac{\partial L}{\partial \ddot{q}_i}\right) - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) + \frac{\partial L}{\partial q_i} = 0 \ .$$

Estenda este resultado para lagrangianas contendo derivadas de terceira ordem das coordenadas generalizadas. (ii) Mostre que a lagrangiana $L = -mq\ddot{q}/2 - kq^2/2$ gera a equação de movimento de um oscilador harmônico. Prove que esta lagrangiana difere

da lagrangiana usual de um oscilador harmônico por uma derivada total em relação ao tempo.

2.14. No caso de amortecimento fraco, mostre que com $\Omega = (\omega^2 - \lambda^2/4)^{1/2}$ a lagrangiana

$$L = \frac{2\dot{x} + \lambda x}{2\Omega x} \tan^{-1}\left(\frac{2\dot{x} + \lambda x}{2\Omega x}\right) - \frac{1}{2}\ln(\dot{x}^2 + \lambda x\dot{x} + \omega^2 x^2)$$

gera a equação de movimento de um oscilador harmônico amortecido (Havas 1957). Prove que a ação associada a esta lagrangiana é quase-invariante sob a dilatação espacial infinitesimal $x'(t') = (1+\epsilon)x(t)$, $t' = t$, isto é, que a Eq.(2.7.11) é verdadeira com $G(x,t) = -t$. Aplique, em seguida, o teorema de Noether generalizado para deduzir que

$$-\tan^{-1}\left(\frac{2\dot{x} + \lambda x}{2\Omega x}\right) - \Omega t = \delta \ ,$$

onde δ é uma constante. Conjugue este resultado com a constante de movimento (2.7.17) para obter por meios puramente algébricos (Lemos 1993)

$$x(t) = Ae^{-\lambda t/2}\cos(\Omega t + \delta) \ ,$$

que é a solução geral da equação de movimento do oscilador amortecido.

2.15. Considere um sistema de N partículas em interação com lagrangiana $L = T - V$, onde $T = \sum_i m_i \dot{\mathbf{r}}_i^2/2$ e V só depende das distâncias $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ entre as partículas. Mostre que, sob uma transformação de Galileu $\mathbf{r}_i{}' = \mathbf{r}_i - \mathbf{v}t$, $t' = t$, a lagrangiana transformada só difere da lagrangiana original pela derivada total de uma certa função $F(\mathbf{r}_1, \ldots, \mathbf{r}_N, t)$. Tomando $\mathbf{v}$ infinitesimal, mostre que a ação é quase invariante e, como conseqüência da versão generalizada do teorema de Noether, obtenha a constante de movimento $\mathbf{C} = t\mathbf{P} - M\mathbf{R}$, onde $\mathbf{P}$ é o momento linear total, M é a massa total e $\mathbf{R}$ é o vetor posição do centro de massa do sistema.

2.16. Uma partícula de massa unitária move-se ao longo de uma linha reta sujeita à força não-conservativa $F = -\gamma\dot{x}^2$, com γ constante. (i) Mostre que a equação de movimento da partícula pode ser obtida da lagrangiana $L = e^{2\gamma x}\dot{x}^2/2$. (ii) Mostre que a integral de Jacobi h associada a esta lagrangiana é equivalente a $\dot{x}e^{\gamma x} = C_1$, onde C_1 é uma constante. (iii) Prove que a ação é invariante sob a transformação $x' = x + a$, $t' = e^{\beta a}t$ para um certo valor de β (determine-o). Interprete geometricamente o significado desta transformação. (iv) Considerando a versão infinitesimal da transformação do item anterior, use o teorema de Noether para provar que $(\gamma t\dot{x} - 1)\dot{x}e^{2\gamma x} = C_2 =$ constante. (v) Combinando as duas constantes de movimento para eliminar $\dot{x}$, prove que

$$x(t) = A + \frac{1}{\gamma}\ln(B + \gamma t)$$

e verifique diretamente que esta é a solução geral da equação de movimento da partícula.

2.17. Um sistema com dois graus de liberdade tem lagrangiana

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - (\alpha x + \beta y) \ ,$$

com α e β constantes. (i) Prove que a lagrangiana – e, conseqüentemente, a ação – é invariante sob a transformação infinitesimal $x \to x' = x + \epsilon\beta$, $y \to y' = y - \epsilon\alpha$. Usando o teorema de Noether, prove que a quantidade A definida por

$$A = m(\beta\dot{x} - \alpha\dot{y})$$

é constante de movimento. (ii) Introduzindo as novas coordenadas generalizadas $\bar{x}$ e $\bar{y}$ definidas por $\bar{x} = \alpha x + \beta y$ e $\bar{y} = \beta x - \alpha y$, mostre que uma das novas coordenadas não aparece na lagrangiana expressa em termos de $\bar{x}, \bar{y}, \dot{\bar{x}}, \dot{\bar{y}}$. Prove que a quantidade A do item (i) é proporcional ao momento conjugado a essa coordenada cíclica e interprete geometricamente todo o procedimento realizado.

2.18. Uma partícula de massa m e carga elétrica e move-se num campo magnético uniforme $\mathbf{B} = B\hat{\mathbf{k}}$, cujo potencial vetor é $\mathbf{A}(\mathbf{r}) = \mathbf{B} \times \mathbf{r}/2$. (i) Mostre que as equações de Lagrange para a partícula equivalem a $\dot{\mathbf{v}} = -\boldsymbol{\omega} \times \mathbf{v}$ e determine $\boldsymbol{\omega}$ em termos de $\mathbf{B}$. (ii) Exprima a lagrangiana em coordenadas cilíndricas (ρ, φ, z) e mostre que, embora φ seja coordenada cíclica, a componente $L_z = m\rho^2\dot{\varphi}$ do momento angular *não é conservada.* Explique.

2.19. Seja q_k uma coordenada cíclica da lagrangiana $L(q, \dot{q}, t)$, de modo que o momento conjugado p_k é constante de movimento. Se L for substituída por $\bar{L} = L + dF(q,t)/dt$, em geral q_k não será coordenada cíclica de $\bar{L}$ e o momento conjugado não se conservará. Mas sabemos que $\bar{L}$ e L produzem as mesmas equações de movimento. Resolva o aparente paradoxo: como é possível que as mesmas equações de movimento impliquem tanto conservação quanto não conservação? Escolhendo $F = q^2$, discuta o caso da partícula livre unidimensional com $L = \dot{q}^2/2$.

2.20. Reconsidere o Problema 1.12 pela técnica dos multiplicadores de Lagrange e determine a força de vínculo sobre o bloco de massa m.

2.21. Mostre que a lagrangiana do Problema 1.7 é quase-invariante sob uma translação infinitesimal $x' = x + \epsilon$ e encontre a constante de movimento correspondente via teorema de Noether. Usando a equação de movimento do sistema, verifique diretamente que a quantidade obtida pelo teorema de Noether é de fato constante de movimento.

2.22. Uma partícula de massa m e vetor posição $\mathbf{r}$ move-se no potencial coulombiano ou gravitacional $V = -\kappa/r$, onde $r = |\mathbf{r}|$. (i) Mostre que o vetor de Laplace-Runge-Lenz

$$\mathbf{A} = \mathbf{p} \times \boldsymbol{l} - m\kappa\frac{\mathbf{r}}{r}$$

é constante de movimento, onde $\boldsymbol{l} = \mathbf{r} \times \mathbf{p}$ é o vetor momento angular. (ii) Mostre que $\mathbf{A}$ está contido no plano da órbita. (iii) Tomando o produto escalar de $\mathbf{A}$ por $\mathbf{r}$, deduza a equação da órbita. Sugestão: $\mathbf{r} \cdot (\mathbf{p} \times \boldsymbol{l}) = \boldsymbol{l} \cdot \boldsymbol{l}$. (iv) Mostre que a excentricidade é dada por $e = A/m\kappa$ e que $\mathbf{A}$ aponta ao longo do eixo maior da órbita elíptica.

2.23. Considere a lagrangiana unidimensional $L = \dot{x}^2/2 - g/x^2$, com g constante. (i) Mostre que a ação é invariante sob a transformação finita $x'(t') = e^{\alpha}\, x(t)$, $t' = e^{2\alpha}\, t$, onde α é uma constante arbitrária, e conclua, via teorema de Noether, que $I = x\dot{x} - 2Et$ é constante de movimento, com E denotando a energia total. (ii) Prove que a ação é quase-invariante sob a transformação infinitesimal $t' = t + \epsilon t^2$, $x'(t') = x(t) - \epsilon t x(t)$ e

daí conclua que $K = Et^2 - tx\dot{x} + x^2/2$ é constante de movimento. (iii) Combinando os resultados dos itens (i) e (ii) obtenha a solução geral da equação de Lagrange para $x(t)$.

2.24. Mostre que, na presença vínculos da forma (2.4.10), a Eq. (2.6.2) deve ser substituída por

$$\frac{dh}{dt} = -\frac{\partial L}{\partial t} - \sum_{l=1}^{p} \lambda_l a_{lt} \ .$$

Assim, mesmo quando L não depende explicitamente do tempo, h em geral não se conserva se os vínculos dependem do tempo.

2.25. Reconsidere o problema de um cilindro rolando sem deslizar sobre outro cilindro fixo, Exemplo 2.4.3. (i) Mostre que a força de vínculo tangencial sobre o cilindro móvel é $f_\theta = -(mg/3)\,\mathrm{sen}\,\theta$. Esta é a força de atrito necessária para manter o rolamento sem deslizamento. (ii) Se μ é o coeficiente de atrito estático entre os cilindros, só não ocorrerá deslizamento enquanto $|f_\theta| \leq \mu N$, onde N é a força de reação normal dada por (2.4.49). Mostre que o deslizamento começa quando o ângulo θ atinge o valor θ_1 tal que

$$\cos\theta_1 = \frac{28\mu^2 + \sqrt{1+33\mu^2}}{1+49\mu^2} \ .$$

(iii) Mostre que o resultado do Exemplo 2.4.3 para o ângulo em que o cilindro móvel abandona o cilindro fixo só é correto se $\mu \gg 1$.

2.26 Uma partícula de carga e move-se num campo eletromagnético uniforme, isto é, $\mathbf{E}$ e $\mathbf{B}$ são constantes. Mostre que o vetor $\mathbf{K} = m\mathbf{v} + \frac{e}{c}\mathbf{B} \times \mathbf{r} - e\mathbf{E}\,t$ é constante de movimento.

Capítulo 3

CINEMÁTICA DA ROTAÇÃO

> For a physicist mathematics is not just a tool by means of which phenomena can be calculated, it is the main source of concepts and principles by means of which new theories can be created.
>
> Freeman Dyson

A dinâmica do corpo rígido é um capítulo da mecânica clássica que merece destaque não apenas por seu interesse intrínseco, mas também porque requer a introdução de conceitos físicos e técnicas matemáticas sumamente importantes. Antes, entretanto, de enveredar pela dinâmica, é preciso introduzir métodos eficazes para descrever o movimento de corpos rígidos. Dedicaremos um espaço considerável ao estudo da cinemática da rotação na perspectiva de que diversos dos instrumentos matemáticos a serem desenvolvidos são de grande generalidade, encontrando ampla aplicação em outros domínios da física teórica.

3.1 Transformações Ortogonais

Um corpo rígido possui, em geral, seis graus de liberdade. Obviamente, três desses graus de liberdade correspondem às translações do corpo como um todo, ao passo que os outros três graus de liberdade descrevem as diferentes orientações do corpo relativamente a um sistema de eixos fixos no espaço. Um modo simples de especificar a orientação do corpo rígido consiste em estabelecer um sistema cartesiano *fixo no corpo*, isto é, que acompanha os seus movimentos, e considerar os ângulos que os eixos atados ao corpo formam com eixos paralelos aos que permanecem fixos no espaço (representados por linhas tracejadas na Fig. 3.1).

Cossenos Diretores

Seja Σ um sistema de coordenadas cartesianas (x_1, x_2, x_3), com os correspondentes vetores unitários $(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3)$, representando eixos fixos no espaço, e seja Σ' um sistema de coordenadas cartesianas (x'_1, x'_2, x'_3), cujos eixos permanecem solidários ao corpo rígido, com vetores unitários $(\hat{\mathbf{e}}'_1, \hat{\mathbf{e}}'_2, \hat{\mathbf{e}}'_3)$, conforme a Fig. 3.2. Um vetor arbitrário $\boldsymbol{g}$ pode ser decomposto tanto na base de Σ quanto na de Σ':

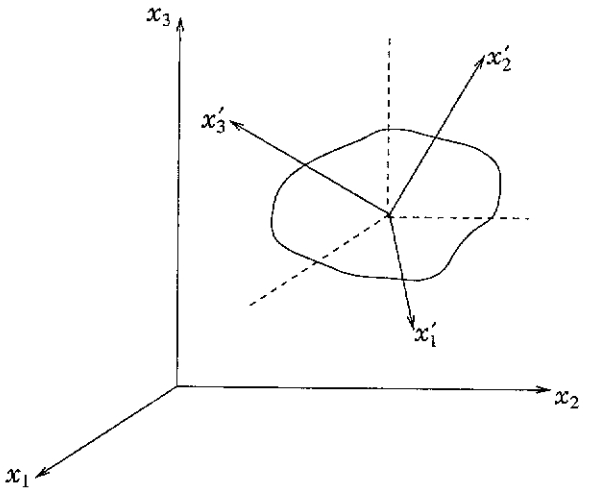


Figura 3.1: Sistemas de eixos fixos no espaço e solidários a um corpo rígido.

$$\boldsymbol{g} = g_1\hat{\mathbf{e}}_1 + g_2\hat{\mathbf{e}}_2 + g_3\hat{\mathbf{e}}_3 = \sum_{j=1}^{3} g_j\hat{\mathbf{e}}_j \ , \tag{3.1.1a}$$

$$\boldsymbol{g} = g_1'\hat{\mathbf{e}}_1' + g_2'\hat{\mathbf{e}}_2' + g_3'\hat{\mathbf{e}}_3' = \sum_{j=1}^{3} g_j'\hat{\mathbf{e}}_j' \ . \tag{3.1.1b}$$

Naturalmente, as componentes de $\boldsymbol{g}$ relativamente a Σ e Σ' estão relacionadas entre si. Com efeito,

$$g_i' = \hat{\mathbf{e}}_i' \cdot \boldsymbol{g} = \sum_{j=1}^{3} g_j \hat{\mathbf{e}}_i' \cdot \hat{\mathbf{e}}_j \equiv \sum_{j=1}^{3} a_{ij}\, g_j \quad , \quad i = 1,2,3 \ , \tag{3.1.2}$$

onde

$$a_{ij} = \hat{\mathbf{e}}_i' \cdot \hat{\mathbf{e}}_j = \text{cosseno do ângulo entre os eixos } x_i' \text{ e } x_j \ . \tag{3.1.3}$$

As quantidades a_{ij} — nove ao todo — são chamadas de cossenos diretores de Σ' relativamente a Σ. Portanto, as componentes de $\boldsymbol{g}$ em Σ' exprimem-se em termos de suas componentes em Σ por meio de uma *transformação linear* cujos coeficientes são os cossenos diretores.

As nove grandezas a_{ij} não podem ser mutuamente independentes, pois bastam três quantidades para definir a orientação de Σ' relativamente a Σ. Para obter as condições a que os cossenos diretores devem estar sujeitos, basta levar em conta que o módulo de um vetor não depende da particular base em que ele é decomposto:

$$\boldsymbol{g} \cdot \boldsymbol{g} = \sum_{i=1}^{3} g_i' g_i' = \sum_{i=1}^{3} g_i\, g_i \ . \tag{3.1.4}$$

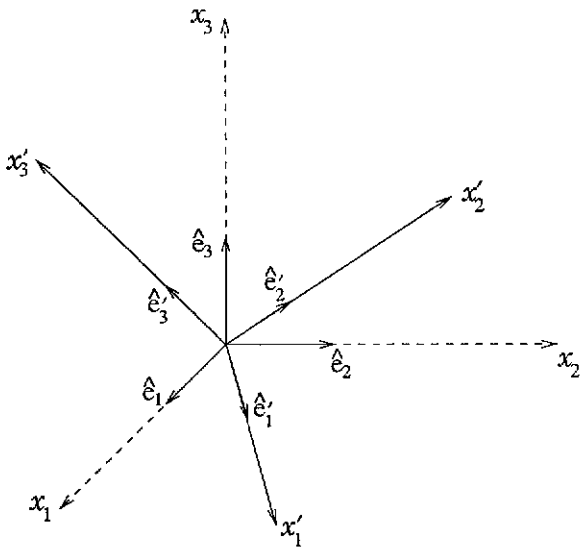


Figura 3.2: Vetores unitários fixos no espaço $(\hat{e}_1, \hat{e}_2, \hat{e}_3)$ e atados ao corpo rígido $(\hat{e}'_1, \hat{e}'_2, \hat{e}'_3)$.

Mas, fazendo uso da Eq.(3.1.2),

$$\sum_{i=1}^{3} g'_i\, g'_i = \sum_{i=1}^{3}(\sum_{j=1}^{3} a_{ij}\, g_j)(\sum_{k=1}^{3} a_{ik}\, g_k) = \sum_{k,j=1}^{3}(\sum_{i=1}^{3} a_{ij} a_{ik}) g_j\, g_k \ . \tag{3.1.5}$$

Utilizando o símbolo delta de Kronecker δ_{jk} introduzido no Apêndice A, é possível escrever

$$\sum_{i=1}^{3} g_i g_i \equiv g_1^2 + g_2^2 + g_3^2 = \sum_{k,j=1}^{3} \delta_{jk} g_j g_k \ . \tag{3.1.6}$$

Comparando as equações (3.1.5) e (3.1.6), levando em consideração (3.1.4) e a arbitrariedade de g_1, g_2, g_3, deduz-se

$$\sum_{i=1}^{3} a_{ij} a_{ik} = \delta_{jk} \ . \tag{3.1.7}$$

Estas são as condições obedecidas pelos nove cossenos diretores, mostrando que nem todos eles são independentes.

Matrizes Ortogonais

A equação (3.1.2) constitui uma representação compacta das três equações

$$g'_1 = a_{11} g_1 + a_{12} g_2 + a_{13} g_3 \ ,$$

$$g'_2 = a_{21}g_1 + a_{22}g_2 + a_{23}g_3 \ , \tag{3.1.8}$$
$$g'_3 = a_{31}g_1 + a_{32}g_2 + a_{33}g_3 \ .$$

Equações deste tipo podem ser tornadas muito mais concisas por meio da notação matricial. Definindo as matrizes[1]

$$\boldsymbol{g}_\Sigma = \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} \ , \ \boldsymbol{g}_{\Sigma'} = \begin{pmatrix} g'_1 \\ g'_2 \\ g'_3 \end{pmatrix} \ , \ \boldsymbol{\mathcal{A}} = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \ , \tag{3.1.9}$$

as equações (3.1.8) podem expressar-se na forma

$$\boldsymbol{g}_{\Sigma'} = \boldsymbol{\mathcal{A}}\boldsymbol{g}_\Sigma \ . \tag{3.1.10}$$

Por outro lado, notando que $a_{ij} = (\boldsymbol{\mathcal{A}}^T)_{ji}$, onde $\boldsymbol{\mathcal{A}}^T$ denota a transposta de $\boldsymbol{\mathcal{A}}$, a Eq.(3.1.7) é equivalente a $\sum_{i=1}^{3}(\boldsymbol{\mathcal{A}}^T)_{ji}(\boldsymbol{\mathcal{A}})_{ik} = \delta_{jk}$. Assim, considerando a definição de produto de matrizes, a Eq.(3.1.7) pode ser reescrita na forma matricial

$$\boldsymbol{\mathcal{A}}^T\boldsymbol{\mathcal{A}} = \boldsymbol{I} \ , \tag{3.1.11}$$

onde

$$\boldsymbol{I} = (\delta_{jk}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.1.12}$$

é a matriz identidade 3×3.

■ **Exercício 3.1.1.** Prove que a transformação inversa de (3.1.2) é

$$g_i = \sum_{j=1}^{3} a_{ji}\, g'_j \tag{3.1.13}$$

e, usando a invariância de $\boldsymbol{g}.\boldsymbol{g}$, prove que

$$\sum_{i=1}^{3} a_{ji}\, a_{ki} = \delta_{jk} \ . \tag{3.1.14}$$

Mostre que, em notação matricial, esta equação toma a forma[2]

$$\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{A}}^T = \boldsymbol{I} \ . \tag{3.1.15}$$

[1] Neste capítulo usaremos letras maiúsculas caligráficas em negrito, tais como $\boldsymbol{\mathcal{A}}$, para representar matrizes quadradas. Um vetor escrito na forma de matriz-coluna será representado por uma letra minúscula em negrito. Quando necessário, será acrescentado um subscrito indicando a base em que o vetor está expresso, tal como em $\boldsymbol{g}_\Sigma$.

[2] Para matrizes finitas as Eq.(3.1.11) e (3.1.15) são equivalentes, isto é, uma é verdadeira se e somente se a outra é verdadeira (vide o Apêndice C). Esta propriedade não se estende a matrizes infinitas.

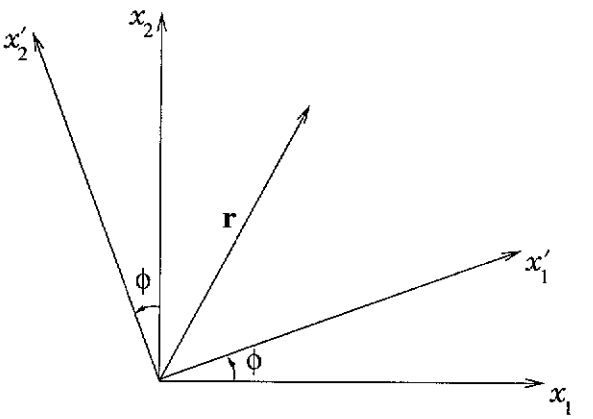


Figura 3.3: Rotação num plano.

Demonstre, ainda, as leis de transformação direta e inversa

$$\hat{\mathbf{e}}_j' = \sum_{l=1}^{3} a_{jl}\,\hat{\mathbf{e}}_l \qquad , \qquad \hat{\mathbf{e}}_j = \sum_{l=1}^{3} a_{lj}\,\hat{\mathbf{e}}_l' \ . \tag{3.1.16}$$

As bases $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$ e $\{\hat{\mathbf{e}}_1', \hat{\mathbf{e}}_2', \hat{\mathbf{e}}_3'\}$ são compostas por vetores unitários e mutuamente ortogonais, logo valem as equações

$$\hat{\mathbf{e}}_j \cdot \hat{\mathbf{e}}_k = \hat{\mathbf{e}}_j' \cdot \hat{\mathbf{e}}_k' = \delta_{jk} \ . \tag{3.1.17}$$

Deduza, finalmente, (3.1.7) e (3.1.14) a partir de (3.1.17). ■

A matriz de transformação entre sistemas cartesianos obedece às equações (3.1.11) e (3.1.15), as quais são equivalentes a

$$\mathcal{A}^{-1} = \mathcal{A}^T \ . \tag{3.1.18}$$

Qualquer matriz que obedeça a esta última equação é dita uma *matriz ortogonal*, e a transformação linear associada (3.1.10) é chamada de *transformação ortogonal*.

■ **Exemplo 3.1.1** No caso de rotações no plano (Fig. 3.3), temos

$$x_1' = x_1 \cos\phi + x_2 \operatorname{sen}\phi \ , \tag{3.1.19a}$$

$$x_2' = -x_1 \operatorname{sen}\phi + x_2 \cos\phi \ . \tag{3.1.19b}$$

A matriz associada a esta transformação ortogonal é

$$\mathcal{A} = \begin{pmatrix} \cos\phi & \operatorname{sen}\phi \\ -\operatorname{sen}\phi & \cos\phi \end{pmatrix} \ , \tag{3.1.20}$$

e um único parâmetro – o ângulo de rotação ϕ – especifica completamente a transformação.■

■ **Exercício 3.1.2.** Usando a definição (3.1.3) e a equação (3.1.2) com $g = \mathrm{r}$, demonstre (3.1.19). Verifique que a matriz $\mathcal{A}$ dada por (3.1.20) é ortogonal. ■

A matriz de transformação $\mathcal{A}$ pode ser encarada como uma transformação linear que, conforme a Eq.(3.1.10), age sobre as componentes de um vetor em Σ e as transforma nas componentes do *mesmo vetor* em Σ'. Esta interpretação da Eq.(3.1.10), em que o vetor permanece o mesmo e o que muda é o sistema de coordenadas, é denominada *ponto de vista passivo*. Alternativamente, $\mathcal{A}$ pode ser encarado como um operador que transforma o vetor $\boldsymbol{g}$ num *outro vetor* $\boldsymbol{g}'$, ambos considerados expressos no mesmo sistema de coordenadas. Esta forma de interpretar a transformação executada pela matriz $\mathcal{A}$ é conhecida como *ponto de vista ativo*, e em muitas ocasiões, como veremos, revela-se mais conveniente do que o ponto de vista passivo. Naturalmente, a operação específica representada por $\mathcal{A}$ dependerá da interpretação adotada. Assim, por exemplo, uma rotação dos eixos coordenados mantendo o vetor fixo faz o mesmo efeito que manter os eixos fixos e girar o vetor do mesmo ângulo, porém no sentido contrário. No Exemplo 3.1.1, se $\mathcal{A}$ representa uma rotação anti-horária de um ângulo ϕ dos eixos coordenados, ela corresponderá a uma rotação horária do mesmo ângulo ϕ aplicada ao vetor, e vice-versa.

Grupo das Matrizes Ortogonais

É necessário examinar transformações ortogonais sucessivas. Seja $\mathcal{B} = (b_{ij})$ a matriz de transformação de Σ para Σ', e $\mathcal{A} = (a_{ij})$ a matriz de transformação de Σ' para Σ'', de modo que

$$g'_i = \sum_j b_{ij}\, g_j \ , \tag{3.1.21}$$

$$g''_k = \sum_i a_{ki}\, g'_i \ . \tag{3.1.22}$$

Conseqüentemente,

$$g''_k = \sum_{ij} a_{ki}\, b_{ij}\, g_j = \sum_j (\sum_i a_{ki}\, b_{ij})\, g_j \equiv \sum_j c_{kj}\, g_j \ , \tag{3.1.23}$$

onde os c_{kj} definidos por

$$c_{kj} = \sum_i a_{ki}\, b_{ij} \tag{3.1.24}$$

são os elementos da matriz $\mathcal{C}$ que realiza a transformação direta de Σ para Σ''. De acordo com a definição de produto de matrizes, a Eq.(3.1.24) é equivalente a

$$\mathcal{C} = \mathcal{A}\mathcal{B} \ . \tag{3.1.25}$$

É imprescindível observar a ordem dos fatores nesta última equação, uma vez que o produto de matrizes não é comutativo: a matriz associada à primeira transformação aparece na posição mais à direita no produto, imediatamente à esquerda vem a matriz

da segunda transformação, e assim por diante. É intuitivamente claro que duas rotações sucessivas são equivalentes a uma única rotação, de sorte que, se $\mathcal{A}$ e $\mathcal{B}$ são matrizes ortogonais, seu produto também tem que ser ortogonal. Para provar que, como indica nossa intuição, $\mathcal{C}$ é ortogonal, basta mostrar que $\mathcal{C}$ satisfaz (3.1.15), por exemplo. Como primeiro passo, note que

$$\mathcal{C}\mathcal{C}^T = \mathcal{A}\mathcal{B}(\mathcal{A}\mathcal{B})^T = \mathcal{A}\mathcal{B}\mathcal{B}^T\mathcal{A}^T \ ,$$

onde usamos a identidade $(\mathcal{A}\mathcal{B})^T = \mathcal{B}^T\mathcal{A}^T$. Mas, com o uso da ortogonalidade de $\mathcal{B}$ e de $\mathcal{A}$, podemos escrever

$$\mathcal{C}\mathcal{C}^T = \mathcal{A}\mathcal{I}\mathcal{A}^T = \mathcal{A}\mathcal{A}^T = \mathcal{I} \ ,$$

o que comprova a ortogonalidade de $\mathcal{C}$.

O conjunto das matrizes ortogonais goza das seguintes propriedades:

(1) Se $\mathcal{A}$ e $\mathcal{B}$ são elementos do conjunto, o produto $\mathcal{A}\mathcal{B}$ também o é.
(2) O produto é *associativo*, isto é, $\mathcal{A}(\mathcal{B}\mathcal{C}) = (\mathcal{A}\mathcal{B})\mathcal{C}$ para quaisquer elementos $\mathcal{A}, \mathcal{B}, \mathcal{C}$ do conjunto.
(3) O conjunto contém um elemento $\mathcal{I}$, chamado *identidade*, tal que, para qualquer elemento $\mathcal{A}$ do conjunto, $\mathcal{A}\mathcal{I} = \mathcal{I}\mathcal{A} = \mathcal{A}$.
(4) Para cada elemento $\mathcal{A}$ do conjunto existe um único elemento $\mathcal{B}$ do conjunto tal que $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A} = \mathcal{I}$. O elemento $\mathcal{B}$ é denominado *inverso* de $\mathcal{A}$ e denotado por $\mathcal{A}^{-1}$.

A primeira propriedade já foi demonstrada, a segunda é satisfeita porque o produto de matrizes é associativo, e a terceira é obviamente verdadeira. A quarta propriedade é a única que ainda requer prova. Mas para qualquer matriz ortogonal a única inversa simplesmente coincide com a transposta, e verifica-se imediatamente que a transposta de uma matriz ortogonal é também uma matriz ortogonal porque podemos escrever $(\mathcal{A}^T)^T\mathcal{A}^T = \mathcal{A}\mathcal{A}^T = \mathcal{I}$.

Um conjunto de elementos no qual está definida uma *lei de composição* — também chamada de *multiplicação* ou *produto* — com as quatro propriedades acima constitui um *grupo* (Hamermesh 1962; Tung 1985). O grupo das matrizes ortogonais 3×3 é denotado por $O(3)$. A teoria dos grupos desempenha um papel importantíssimo nos mais diversos ramos da física teórica atual, principalmente na física das partículas elementares.

3.2 Deslocamentos Possíveis de um Corpo Rígido

O determinante de uma matriz ortogonal só pode assumir certos valores restritos. De fato, tomando o determinante de (3.1.15) resulta

$$(\det \mathcal{A})(\det \mathcal{A}^T) = 1 \quad \Longrightarrow \quad (\det \mathcal{A})^2 = 1 \ , \tag{3.2.1}$$

onde utilizamos as seguintes e bem conhecidas propriedades dos determinantes: $\det(\mathcal{A}\mathcal{B}) = (\det \mathcal{A})(\det \mathcal{B})$; $\det \mathcal{A}^T = \det \mathcal{A}$. Portanto, uma matriz ortogonal só pode ter determinante igual a $+1$ ou -1. Uma matriz ortogonal simples com determinante -1 é

$$\mathcal{I} = -\boldsymbol{I} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} . \tag{3.2.2}$$

A transformação efetuada por $\mathcal{I}$ é

$$g_1' = -g_1 \quad , \quad g_2' = -g_2 \quad , \quad g_3' = -g_3 \ , \tag{3.2.3}$$

que corresponde a uma inversão dos eixos coordenados.

Uma inversão transforma um sistema cartesiano orientado segundo a mão direita num sistema cartesiano orientado segundo a mão esquerda. É claro, portanto, que uma inversão não corresponde a um deslocamento possível de um corpo rígido: uma vez escolhido um sistema cartesiano de mão direita fixo no corpo rígido, nenhum deslocamento físico do corpo pode alterar a orientação do referido sistema cartesiano, salvo se o corpo for deformável, o que não é o caso por hipótese. Por outro lado, qualquer matriz ortogonal com determinante -1 pode ser escrita na forma $\mathcal{A} = (-\boldsymbol{I})(-\mathcal{A}) \equiv \mathcal{I}\mathcal{B}$ onde $\det \mathcal{B} \equiv \det(-\mathcal{A}) = +1$, isto é, $\mathcal{A}$ contém uma inversão dos eixos e não representa um deslocamento possível de um corpo rígido. Matrizes ortogonais com determinante $+1$ são ditas *próprias*, ao passo que aquelas com determinante -1 são ditas *impróprias*. Concluímos, assim, que somente as transformações ortogonais próprias correspondem a deslocamentos possíveis de um corpo rígido. Demonstra-se facilmente que o conjunto das transformações ortogonais próprias constitui um grupo, chamado de *grupo de rotações* e denotado por $SO(3)$.

Um resultado clássico que revela uma das características fundamentais do movimento de um corpo rígido foi estabelecido por Euler no século XVIII, mais precisamente em 1776.

Teorema 3.2.1 (Euler). O deslocamento mais geral de um corpo rígido com um ponto fixo é uma rotação em torno de algum eixo.

Demonstração. Em vez da prova geométrica tradicional, recorreremos às técnicas algébricas previamente desenvolvidas, o que conferirá à demonstração um caráter bastante instrutivo. Antes, porém, é necessário esclarecer melhor o conteúdo do teorema. Suponha que o corpo rígido esteja numa configuração inicial e, após mover-se de forma arbitrária em torno de um ponto fixo, atinja uma configuração final. Como já discutimos, a configuração final resulta da aplicação de uma transformação ortogonal própria à configuração inicial. Segundo o enunciado do teorema de Euler, a configuração final pode ser obtida por meio de uma única rotação em torno de algum eixo contendo o ponto fixo, isto é, uma rotação em torno de um ponto é sempre equivalente a uma rotação em torno de um eixo passando pelo ponto. Mas um vetor ao longo do eixo de rotação fica inalterado, suas componentes antes e depois da rotação permanecem as mesmas. Se $\mathcal{A}$ é a matriz ortogonal própria que leva o corpo da sua configuração inicial à final, o teorema de Euler estará demonstrado desde que se prove a existência de um vetor não-nulo $\boldsymbol{n}$ tal que

$$\mathcal{A}\boldsymbol{n} = \boldsymbol{n} \tag{3.2.4}$$

ou, equivalentemente,

$$(\mathcal{A} - \boldsymbol{I})\,\boldsymbol{n} = 0 \ . \tag{3.2.5}$$

Se a matriz $(\mathcal{A} - \boldsymbol{I})$ for inversível, multiplicando-se (3.2.5) pela esquerda por $(\mathcal{A} - \boldsymbol{I})^{-1}$ deduz-se que $\boldsymbol{n} = 0$. Conseqüentemente, existirá $\boldsymbol{n} \neq 0$ satisfazendo (3.2.5) se e somente se $(\mathcal{A} - \boldsymbol{I})$ não possuir inversa, isto é, se e somente se

$$\det(\mathcal{A} - \boldsymbol{I}) = 0 \ . \tag{3.2.6}$$

Em suma, demonstrar o teorema de Euler equivale a provar a veracidade de (3.2.6). A partir da identidade

$$\mathcal{A} - \boldsymbol{I} = \mathcal{A} - \mathcal{A}\mathcal{A}^T = \mathcal{A}(\boldsymbol{I} - \mathcal{A}^T) = -\mathcal{A}(\mathcal{A} - \boldsymbol{I})^T = (-\boldsymbol{I})\mathcal{A}(\mathcal{A} - \boldsymbol{I})^T \tag{3.2.7}$$

deduz-se

$$\det(\mathcal{A} - \boldsymbol{I}) = \det(-\boldsymbol{I}) \det \mathcal{A} \det(\mathcal{A} - \boldsymbol{I}) = -\det \mathcal{A} \det(\mathcal{A} - \boldsymbol{I}) \ . \tag{3.2.8}$$

Usando $\det \mathcal{A} = 1$, a Eq.(3.2.8) implica

$$\det(\mathcal{A} - \boldsymbol{I}) = -\det(\mathcal{A} - \boldsymbol{I}) \ , \tag{3.2.9}$$

da qual infere-se (3.2.6), completando a demonstração do teorema. □

A prova geométrica tradicional obscurece o papel da dimensionalidade do espaço, que entra de forma crucial na demonstração acima. Num espaço de dimensão par, $\det(-\boldsymbol{I}) = 1$ e o argumento anterior fracassa. Por exemplo, num espaço bidimensional nenhum vetor do referido espaço é deixado inalterado por uma rotação. De fato, para uma rotação num plano o eixo de rotação é perpendicular ao plano e, em conseqüência, qualquer vetor paralelo ao eixo não pertence ao plano. Em síntese, o teorema de Euler não vale em espaços de dimensão par. O teorema de Euler foi generalizado por Chasles, em 1830, nos seguintes termos.

Teorema 3.2.2 (Chasles). O deslocamento mais geral possível de um corpo rígido é uma translação acompanhada de uma rotação. O eixo de rotação pode ser escolhido de tal modo que a translação seja paralela a este eixo.

A primeira parte do teorema de Chasles é óbvia, pois a remoção do vínculo de mover-se com um ponto fixo introduz três graus de liberdade de translação para o corpo. Uma prova geométrica da segunda parte do teorema de Chasles pode ser encontrada em Whittaker (1944) ou Pars (1965). Uma demonstração moderna, com o emprego de métodos puramente algébricos, está disponível em Corben & Stehle (1962).

3.3 Ângulos de Euler

Numa formulação lagrangiana da dinâmica do corpo rígido, os nove cossenos diretores a_{ij} não constituem as coordenadas mais convenientes para descrever a orientação instantânea de um corpo rígido por não serem independentes entre si. As nove equações

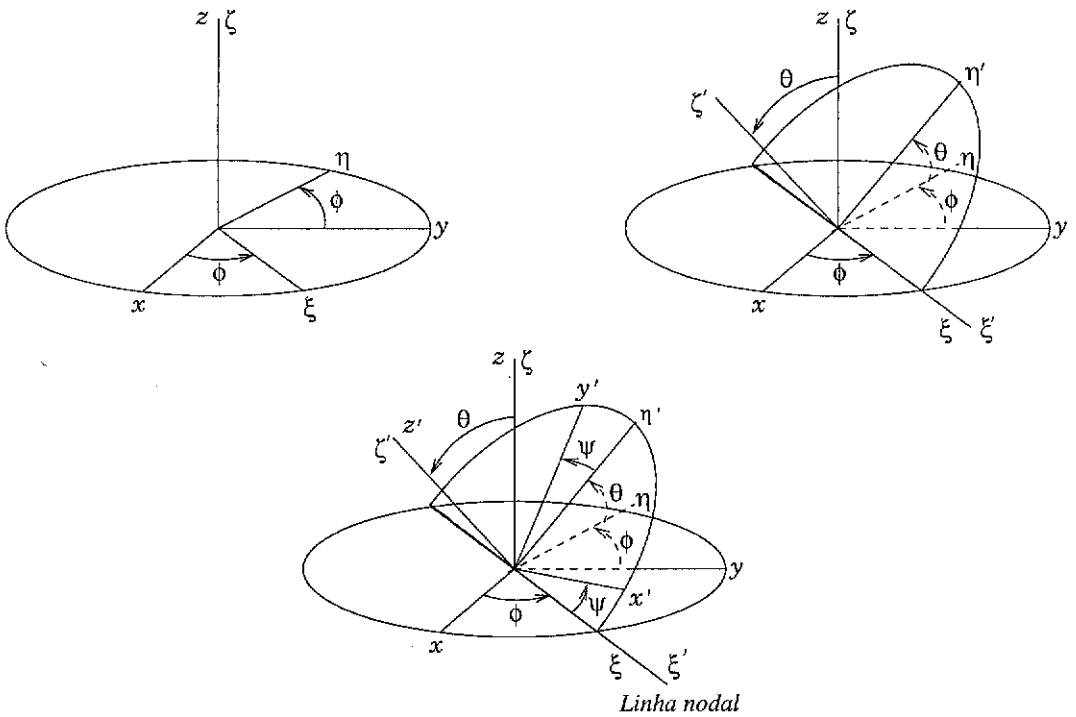


Figura 3.4: Ângulos de Euler.

(3.1.7) impõem apenas seis condições sobre os cossenos diretores, de modo que eles podem ser expressos em função de três parâmetros independentes. Com efeito, há três condições distintas correspondentes à parte diagonal das Eqs.(3.1.7), mas as seis condições correspondentes à parte fora da diagonal ($i \neq j$) são idênticas aos pares. Por exemplo, a equação (3.1.7) com $j = 1, k = 2$ é igual àquela obtida pela escolha $j = 2, k = 1$.

Do ponto de vista prático, uma forma simples e conveniente de parametrizar a matriz de rotação é mediante os ângulos de Euler. A transformação do sistema cartesiano $\Sigma(x, y, z)$ para o sistema $\Sigma'(x', y', z')$ é realizada em três estágios sucessivos, cada um deles servindo para definir um dos ângulos de Euler (Fig. 3.4).

(a) Rotação dos eixos (x, y, z) em torno do eixo z do ângulo ϕ : $(x, y, z) \xrightarrow{\mathcal{D}} (\xi, \eta, \zeta)$.

As equações de transformação confundem-se com as equações (3.1.19) com $x'_1 = \xi$ e $x'_2 = \eta$, acrescidas da equação $x'_3 = \zeta = z$. Logo, a matriz de rotação $\mathcal{D}$ escreve-se

$$\mathcal{D} = \begin{pmatrix} \cos\phi & \operatorname{sen}\phi & 0 \\ -\operatorname{sen}\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{3.3.1}$$

(b) Rotação dos eixos (ξ, η, ζ) em torno do eixo ξ do ângulo θ : $(\xi, \eta, \zeta) \xrightarrow{\mathcal{C}} (\xi', \eta', \zeta')$.

O eixo ξ' é chamado de *linha nodal* e, por analogia com (3.1.19), as equações de

transformação têm a forma

$$\xi' = \xi \ ,$$

$$\eta' = \eta \cos\theta + \zeta \operatorname{sen}\theta \ ,$$

$$\zeta' = -\eta \operatorname{sen}\theta + \zeta \cos\theta \ ,$$

donde

$$\boldsymbol{\mathcal{C}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \operatorname{sen}\theta \\ 0 & -\operatorname{sen}\theta & \cos\theta \end{pmatrix} \ . \tag{3.3.2}$$

(c) Rotação dos eixos (ξ', η', ζ') em torno do eixo ζ' do ângulo ψ : $(\xi', \eta', \zeta') \xrightarrow{\boldsymbol{\mathcal{B}}} (x', y', z')$.

A matriz $\boldsymbol{\mathcal{B}}$ representa uma rotação em torno do terceiro eixo, portanto tem a mesma forma que $\boldsymbol{\mathcal{D}}$:

$$\boldsymbol{\mathcal{B}} = \begin{pmatrix} \cos\psi & \operatorname{sen}\psi & 0 \\ -\operatorname{sen}\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \ . \tag{3.3.3}$$

A rotação $(x, y, z) \rightarrow (x', y', z')$ é executada por $\boldsymbol{\mathcal{A}}_E = \boldsymbol{\mathcal{BCD}}$. Um cálculo direto fornece

$$\boldsymbol{\mathcal{A}}_E = \begin{pmatrix} \cos\psi\cos\phi - \cos\theta\operatorname{sen}\phi\operatorname{sen}\psi & \cos\psi\operatorname{sen}\phi + \cos\theta\cos\phi\operatorname{sen}\psi & \operatorname{sen}\psi\operatorname{sen}\theta \\ -\operatorname{sen}\psi\cos\phi - \cos\theta\operatorname{sen}\phi\cos\psi & -\operatorname{sen}\psi\operatorname{sen}\phi + \cos\theta\cos\phi\cos\psi & \cos\psi\operatorname{sen}\theta \\ \operatorname{sen}\theta\operatorname{sen}\phi & -\operatorname{sen}\theta\cos\phi & \cos\theta \end{pmatrix} \ . \tag{3.3.4}$$

A transformação inversa $(x', y', z') \rightarrow (x, y, z)$ é efetuada por $\boldsymbol{\mathcal{A}}_E^{-1} = \boldsymbol{\mathcal{A}}_E^T$, ou seja,

$$\boldsymbol{\mathcal{A}}_E^{-1} = \begin{pmatrix} \cos\psi\cos\phi - \cos\theta\operatorname{sen}\phi\operatorname{sen}\psi & -\operatorname{sen}\psi\cos\phi - \cos\theta\operatorname{sen}\phi\cos\psi & \operatorname{sen}\theta\operatorname{sen}\phi \\ \cos\psi\operatorname{sen}\phi + \cos\theta\cos\phi\operatorname{sen}\psi & -\operatorname{sen}\psi\operatorname{sen}\phi + \cos\theta\cos\phi\cos\psi & -\operatorname{sen}\theta\cos\phi \\ \operatorname{sen}\psi\operatorname{sen}\theta & \cos\psi\operatorname{sen}\theta & \cos\theta \end{pmatrix} \ . \tag{3.3.5}$$

Em virtude de sua definição, o domínio dos ângulos de Euler é

$$0 \leq \phi \leq 2\pi \quad , \quad 0 \leq \theta \leq \pi \quad , \quad 0 \leq \psi \leq 2\pi \ . \tag{3.3.6}$$

3.4 Rotações Infinitesimais e Velocidade Angular

As equações de movimento de um corpo rígido são equações diferenciais, e sua formulação requer o estudo de rotações do corpo durante um intervalo de tempo infinitesimal.

Comutatividade das Rotações Infinitesimais

Seja $\boldsymbol{g}$ um vetor arbitrário submetido a uma rotação infinitesimal anti-horária de um ângulo $d\Phi$ em torno de um eixo definido pelo vetor unitário $\hat{\boldsymbol{n}}$ (ponto de vista ativo). De acordo com as Eqs.(2.5.10),

$$\boldsymbol{g}' = \boldsymbol{g} + d\boldsymbol{\Omega} \times \boldsymbol{g} \tag{3.4.1}$$

onde

$$d\boldsymbol{\Omega} = d\Phi\, \hat{\boldsymbol{n}} \ . \tag{3.4.2}$$

Os símbolos $d\Phi$ e $d\boldsymbol{\Omega}$ devem ser entendidos como meros nomes para quantidades infinitesimais, e não como diferenciais de um escalar Φ ou de um vetor $\boldsymbol{\Omega}$. Na verdade, em geral não existe nenhum vetor $\boldsymbol{\Omega}$ cuja diferencial seja igual a $d\boldsymbol{\Omega}$ (Problema 3.6).

No caso de rotações sucessivas, com vetores associados $d\boldsymbol{\Omega}_1$ e $d\boldsymbol{\Omega}_2$, escrevemos

$$\boldsymbol{g}' = \boldsymbol{g} + d\boldsymbol{\Omega}_1 \times \boldsymbol{g} \tag{3.4.3}$$

$$\boldsymbol{g}'' = \boldsymbol{g}' + d\boldsymbol{\Omega}_2 \times \boldsymbol{g}' \ , \tag{3.4.4}$$

donde, desprezando infinitésimos de segunda ordem,

$$\boldsymbol{g}'' = \boldsymbol{g} + d\boldsymbol{\Omega}_{12} \times \boldsymbol{g} \ , \tag{3.4.5}$$

com

$$d\boldsymbol{\Omega}_{12} = d\boldsymbol{\Omega}_1 + d\boldsymbol{\Omega}_2 \ . \tag{3.4.6}$$

Este último resultado prova que rotações infinitesimais sucessivas são comutativas ($d\boldsymbol{\Omega}_{12} = d\boldsymbol{\Omega}_{21}$) em decorrência da comutatividade da soma de vetores. Além disso, o vetor associado a rotações infinitesimais sucessivas é a soma dos vetores associados às rotações infinitesimais individuais, propriedade que será de grande valia para os desenvolvimentos subseqüentes.

Vale a pena analisar rotações infinitesimais em linguagem matricial. Adotando o ponto de vista ativo, a equação (3.4.1) pode ser escrita na forma

$$\boldsymbol{g}'_{\Sigma} = \boldsymbol{\mathcal{A}}\, \boldsymbol{g}_{\Sigma} = (\boldsymbol{I} + \boldsymbol{\varepsilon})\, \boldsymbol{g}_{\Sigma} \ , \tag{3.4.7}$$

onde $\boldsymbol{\varepsilon}$ é uma matriz infinitesimal. A comutatividade das rotações infinitesimais é facilmente verificada, já que, com o abandono de infinitésimos de segunda ordem e levando em conta que a soma de matrizes é comutativa,

$$\boldsymbol{\mathcal{A}}_1 \boldsymbol{\mathcal{A}}_2 = (\boldsymbol{I} + \boldsymbol{\varepsilon}_1)(\boldsymbol{I} + \boldsymbol{\varepsilon}_2) = \boldsymbol{I} + \boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2 = \boldsymbol{I} + \boldsymbol{\varepsilon}_2 + \boldsymbol{\varepsilon}_1 = \boldsymbol{\mathcal{A}}_2 \boldsymbol{\mathcal{A}}_1 \ . \tag{3.4.8}$$

Por outro lado, como $\boldsymbol{\mathcal{A}} = \boldsymbol{I} + \boldsymbol{\varepsilon}$ representa uma rotação, $\boldsymbol{\mathcal{A}}$ deve ser matriz ortogonal. Logo,

$$\boldsymbol{I} = \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{A}}^T = (\boldsymbol{I} + \boldsymbol{\varepsilon})(\boldsymbol{I} + \boldsymbol{\varepsilon}^T) = \boldsymbol{I} + \boldsymbol{\varepsilon} + \boldsymbol{\varepsilon}^T \Longrightarrow \boldsymbol{\varepsilon}^T = -\boldsymbol{\varepsilon} \ . \tag{3.4.9}$$

Uma matriz $\boldsymbol{\varepsilon}$ que obedece à equação (3.4.9) é dita anti-simétrica. Em termos dos elementos ϵ_{ij} de $\boldsymbol{\varepsilon}$ a Eq.(3.4.9) equivale a $\epsilon_{ij} = -\epsilon_{ji}$. Portanto, a forma mais geral possível da matriz $\boldsymbol{\varepsilon}$ é

$$\boldsymbol{\varepsilon} = \begin{pmatrix} 0 & \alpha & \beta \\ -\alpha & 0 & \gamma \\ -\beta & -\gamma & 0 \end{pmatrix} \ . \tag{3.4.10}$$

Exigindo que as componentes de (3.4.7) sejam idênticas às de (3.4.1) obtém-se facilmente

$$\boldsymbol{\varepsilon} = \begin{pmatrix} 0 & -d\Omega_3 & d\Omega_2 \\ d\Omega_3 & 0 & -d\Omega_1 \\ -d\Omega_2 & d\Omega_1 & 0 \end{pmatrix} \ . \tag{3.4.11a}$$

Embora, como acabamos de ver, seja possível associar um vetor a cada rotação infinitesimal de tal modo que o vetor associado a duas rotações sucessivas seja a soma dos vetores associados às rotações individuais, tal correspondência não é possível para rotações finitas.[3] Vale ressaltar, ainda, que a Eq.(3.4.11a) estabelece a seguinte correspondência biunívoca entre vetores e matrizes anti-simétricas em três dimensões:

$$\begin{pmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{pmatrix} \quad \longleftrightarrow \quad \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \ . \tag{3.4.11b}$$

Taxa de Variação Temporal de um Vetor

Em geral, a taxa de variação temporal de um vetor medida num sistema de referência solidário a um corpo rígido em rotação não coincidirá com aquela observada de um referencial inercial externo ao corpo. Por exemplo, se o vetor representar a posição de um ponto fixo do corpo rígido, sua taxa de variação temporal será zero em relação ao sistema de eixos fixos no corpo, mas não será nula relativamente ao sistema inercial de eixos em relação ao qual o corpo encontra-se em rotação. Considere um intervalo de tempo infinitesimal dt durante o qual um vetor arbitrário $\boldsymbol{g}$ varia. Intuitivamente, podemos escrever

$$(d\boldsymbol{g})_{inercial} = (d\boldsymbol{g})_{corpo} + (d\boldsymbol{g})_{rot} \ , \tag{3.4.12}$$

pois a diferença entre os dois sistemas deve-se exclusivamente à rotação do corpo. Mas $(d\boldsymbol{g})_{rot}$ é conseqüência de uma rotação infinitesimal de $\boldsymbol{g}$ juntamente com os eixos fixos no corpo, isto é,

$$(d\boldsymbol{g})_{rot} = d\boldsymbol{\Omega} \times \boldsymbol{g} \ , \tag{3.4.13}$$

como se infere da Eq.(3.4.1). A utilização deste último resultado em (3.4.12) conduz a

[3] Um estudo detalhado de rotações finitas pode ser encontrado em Konopinski (1969). Vide, também, o Exemplo 4.2.1 do próximo capítulo.

$$\left(\frac{d\boldsymbol{g}}{dt}\right)_{inercial} = \left(\frac{d\boldsymbol{g}}{dt}\right)_{corpo} + \boldsymbol{\omega} \times \boldsymbol{g} \ , \tag{3.4.14}$$

onde

$$\boldsymbol{\omega} = \frac{d\boldsymbol{\Omega}}{dt} = \frac{d\Phi}{dt}\hat{\boldsymbol{n}} \tag{3.4.15}$$

é a *velocidade angular* do corpo rígido no instante t. Note que o módulo de $\boldsymbol{\omega}$ é a taxa de rotação instantânea, ao passo que sua direção é paralela ao eixo instantâneo de rotação caracterizado pelo vetor unitário $\hat{\boldsymbol{n}}$. Insistimos que a Eq.(3.4.15) representa simbolicamente o limite do quociente de duas grandezas, $d\boldsymbol{\Omega}$ e Δt, que tendem a zero conjuntamente quando $\Delta t \to 0$, e que $\boldsymbol{\omega}$ *não é* a derivada em relação ao tempo de um vetor $\boldsymbol{\Omega}$ (Problema 3.6). Para uma definição puramente analítica de velocidade angular o leitor é remetido ao Problema 3.7.

A importante equação (3.4.14) pode ser deduzida de uma maneira mais rigorosa, que possui a virtude adicional de tornar mais claro o significado de cada um dos seus termos. Sejam $\{\hat{\mathbf{e}}_i\}_{i=1}^3$ os vetores unitários do sistema cartesiano inercial e $\{\hat{\mathbf{e}}'_i\}_{i=1}^3$ os unitários do sistema cartesiano atado ao corpo rígido. Podemos escrever

$$\boldsymbol{g} = \sum_{i=1}^{3} g_i \hat{\mathbf{e}}_i = \sum_{i=1}^{3} g'_i \hat{\mathbf{e}}'_i \ . \tag{3.4.16}$$

Para um observador inercial os $\hat{\mathbf{e}}_i$ são constantes no tempo, mas os $\hat{\mathbf{e}}'_i$ variam com o tempo porque giram juntamente com o corpo. Há, portanto, dois modos distintos de calcular $(d\boldsymbol{g}/dt)_{inercial}$, a saber,

$$\left(\frac{d\boldsymbol{g}}{dt}\right)_{inercial} = \sum_{i=1}^{3} \frac{dg_i}{dt}\, \hat{\mathbf{e}}_i \ , \tag{3.4.17a}$$

e

$$\left(\frac{d\boldsymbol{g}}{dt}\right)_{inercial} = \sum_{i=1}^{3} \frac{dg'_i}{dt}\, \hat{\mathbf{e}}'_i + \sum_{i=1}^{3} g'_i\, \frac{d\hat{\mathbf{e}}'_i}{dt} \ . \tag{3.4.17b}$$

A equação (3.4.17a) define $(d\boldsymbol{g}/dt)_{inercial}$ e, analogamente, $(d\boldsymbol{g}/dt)_{corpo}$ define-se por

$$\left(\frac{d\boldsymbol{g}}{dt}\right)_{corpo} = \sum_{i=1}^{3} \frac{dg'_i}{dt}\, \hat{\mathbf{e}}'_i \ . \tag{3.4.18}$$

Esta última definição é perfeitamente natural visto que, para um observador montado no corpo rígido, os vetores unitários $\hat{\mathbf{e}}'_i$ são constantes. Mas, no intervalo de tempo dt o sistema móvel de eixos sofre uma rotação infinitesimal, o mesmo ocorrendo com cada unitário $\hat{\mathbf{e}}'_i$, isto é,

$$d\hat{\mathbf{e}}'_i = d\boldsymbol{\Omega} \times \hat{\mathbf{e}}'_i \ , \tag{3.4.19}$$

onde usamos (3.4.13) com $\boldsymbol{g} = \hat{\mathbf{e}}'_i$. Substituindo (3.4.18) e (3.4.19) em (3.4.17b) resulta

$$\left(\frac{d\boldsymbol{g}}{dt}\right)_{inercial} = \left(\frac{d\boldsymbol{g}}{dt}\right)_{corpo} + \sum_{i=1}^{3} g'_i \boldsymbol{\omega} \times \hat{\mathbf{e}}'_i = \left(\frac{d\boldsymbol{g}}{dt}\right)_{corpo} + \boldsymbol{\omega} \times \sum_{i=1}^{3} g'_i \hat{\mathbf{e}}'_i \ ,$$

que coincide com (3.4.14).

A igualdade (3.4.14) é válida qualquer que seja o vetor $\boldsymbol{g}$. Isto permite interpretá-la não apenas como uma mera igualdade de vetores, mas como refletindo uma equivalência entre operadores lineares que atuam sobre vetores. Esta igualdade entre operadores pode ser convenientemente expressa na forma simbólica

$$\left(\frac{d}{dt}\right)_{inercial} = \left(\frac{d}{dt}\right)_{corpo} + \boldsymbol{\omega} \times \quad . \tag{3.4.20}$$

■ **Exercício 3.4.1.** Mostre que $(d/dt)_{inercial} = (d/dt)_{corpo}$ quando atuando sobre um escalar. Sugestão: considere o escalar $s = \mathbf{g} \cdot \mathbf{h}$. ■

Unicidade do Vetor Velocidade Angular

Intuitivamente, espera-se que o vetor velocidade angular seja uma propriedade do corpo rígido como um todo, independente, portanto, do ponto escolhido como origem do sistema de coordenadas fixo no corpo. No entanto, uma demonstração rigorosa desse fato é recomendável (Lemos 2000b). Na Fig. 3.5, Σ é um sistema de coordenadas inercial com origem no ponto O, ao passo que Σ'_1 e Σ'_2 são dois sistemas de eixos fixos no corpo, com correspondentes origens O_1 e O_2. Seja $\mathbf{r}$ o vetor posição de um ponto P qualquer do corpo rígido, e sejam $\boldsymbol{\omega}_1$ e $\boldsymbol{\omega}_2$ as velocidades angulares associadas a Σ'_1 e Σ'_2, respectivamente. Uma vez que $\mathbf{r} = \mathbf{R}_1 + \mathbf{r}_1 = \mathbf{R}_2 + \mathbf{r}_2$, temos

$$\left(\frac{d\mathbf{r}}{dt}\right)_{\Sigma} = \left(\frac{d\mathbf{R}_1}{dt}\right)_{\Sigma} + \left(\frac{d\mathbf{r}_1}{dt}\right)_{\Sigma} = \left(\frac{d\mathbf{R}_1}{dt}\right)_{\Sigma} + \boldsymbol{\omega}_1 \times \mathbf{r}_1 \tag{3.4.21}$$

e, analogamente,

$$\left(\frac{d\mathbf{r}}{dt}\right)_{\Sigma} = \left(\frac{d\mathbf{R}_2}{dt}\right)_{\Sigma} + \boldsymbol{\omega}_2 \times \mathbf{r}_2 \ , \tag{3.4.22}$$

porque, sendo P um ponto do corpo, $\mathbf{r}_1$ e $\mathbf{r}_2$ são vetores constantes do ponto de vista dos sistemas de coordenadas Σ'_1 e Σ'_2. Pondo $\mathbf{R} = \mathbf{R}_2 - \mathbf{R}_1$, de (3.4.21) e (3.4.22) deduz-se

$$\left(\frac{d\mathbf{R}}{dt}\right)_{\Sigma} = \boldsymbol{\omega}_1 \times \mathbf{r}_1 - \boldsymbol{\omega}_2 \times \mathbf{r}_2 \ . \tag{3.4.23}$$

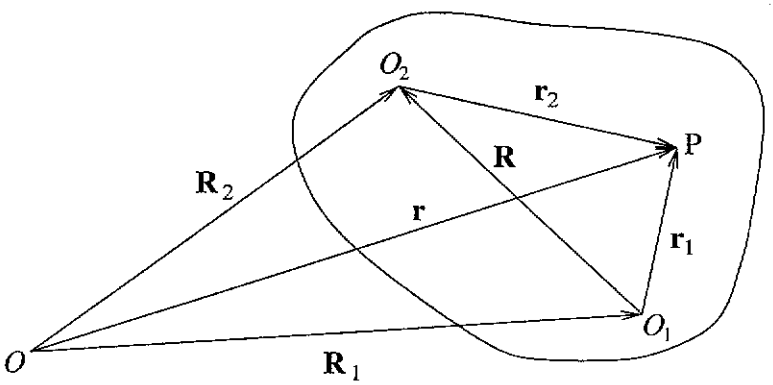


Figura 3.5: O_1 e O_2 são as origens de dois sistemas de coordenadas fixos em pontos distintos de um corpo rígido.

Por outro lado,

$$\left(\frac{d\mathbf{R}}{dt}\right)_{\Sigma} = \left(\frac{d\mathbf{R}}{dt}\right)_{\Sigma_1'} + \omega_1 \times \mathbf{R} = \omega_1 \times \mathbf{R} \ , \tag{3.4.24}$$

pois $\mathbf{R}$ é um vetor constante relativamente ao sistema de coordenadas Σ_1' atado ao corpo. Combinando (3.4.23) e (3.4.24) obtém-se

$$\omega_1 \times \mathbf{R} = \omega_1 \times \mathbf{r}_1 - \omega_2 \times \mathbf{r}_2 \tag{3.4.25}$$

ou, já que $\mathbf{R} = \mathbf{r}_1 - \mathbf{r}_2$,

$$(\omega_2 - \omega_1) \times \mathbf{r}_2 = 0 \ . \tag{3.4.26}$$

Como P é um ponto qualquer do corpo rígido, $\mathbf{r}_2$ é arbitrário e, conseqüentemente, $\omega_2 = \omega_1$, completando a demonstração.

Velocidade Angular em Termos dos Ângulos de Euler

Alguns problemas importantes da dinâmica do corpo rígido exigem que se exprima o vetor velocidade angular em termos dos ângulos de Euler. Uma rotação infinitesimal do corpo rígido é o resultado de três rotações infinitesimais sucessivas cujas velocidades angulares têm, respectivamente, magnitudes $\dot{\phi}$, $\dot{\theta}$, $\dot{\psi}$. Sejam ω_ϕ, ω_θ e ω_ψ os vetores velocidade angular correspondentes. O vetor velocidade angular ω é dado simplesmente por

$$\omega = \omega_\phi + \omega_\theta + \omega_\psi \tag{3.4.27}$$

em virtude da Eq.(3.4.6). As componentes de ω podem ser obtidas tanto ao longo dos

eixos inerciais (x, y, z) quanto dos eixos fixos no corpo (x', y', z'). Por sua maior utilidade, consideraremos este último caso.

Evidentemente, o vetor velocidade angular ω depende linearmente de $\dot{\phi}$, $\dot{\theta}$ e $\dot{\psi}$. Portanto, para encontrar a forma geral de ω podemos fixar um par de ângulos de Euler de cada vez, determinar a velocidade angular associada à variação do terceiro ângulo e depois somar os resultados (Epstein 1982). Fixando θ e ψ, o eixo z, que é fixo no espaço, também torna-se fixo no corpo (ver Fig. 3.4). Logo, z é o eixo de rotação e ω_ϕ é um vetor paralelo ao eixo z, com componente $\dot{\phi}$, de modo que

$$\begin{pmatrix} (\omega_\phi)_{x'} \\ (\omega_\phi)_{y'} \\ (\omega_\phi)_{z'} \end{pmatrix} = \mathcal{A}_E \begin{pmatrix} 0 \\ 0 \\ \dot{\phi} \end{pmatrix} = \begin{pmatrix} \dot{\phi}\,\mathrm{sen}\,\theta\,\mathrm{sen}\,\psi \\ \dot{\phi}\,\mathrm{sen}\,\theta\,\cos\,\psi \\ \dot{\phi}\cos\,\theta \end{pmatrix} . \tag{3.4.28}$$

Fixando agora ϕ e ψ, a linha nodal torna-se fixa no espaço e no corpo, logo é o eixo de rotação. Conseqüentemente, ω_θ é um vetor que só possui a componente $\dot{\theta}$ paralela ao eixo ξ', de modo que

$$\begin{pmatrix} (\omega_\theta)_{x'} \\ (\omega_\theta)_{y'} \\ (\omega_\theta)_{z'} \end{pmatrix} = \mathcal{B} \begin{pmatrix} \dot{\theta} \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \dot{\theta}\,\cos\,\psi \\ -\dot{\theta}\,\mathrm{sen}\,\psi \\ 0 \end{pmatrix} . \tag{3.4.29}$$

Finalmente, com θ e ϕ fixos o eixo z', que é fixo no corpo, torna-se fixo no espaço. Logo, z' é o eixo de rotação e ω_ψ é um vetor com componente $\dot{\psi}$ ao longo do eixo z', não sendo necessário aplicar-lhe qualquer matriz de transformação. Reunindo as componentes correspondentes chega-se ao resultado final

$$\begin{aligned} \omega_{x'} &= \dot{\phi}\,\mathrm{sen}\,\theta\,\mathrm{sen}\,\psi + \dot{\theta}\,\cos\,\psi \ , \\ \omega_{y'} &= \dot{\phi}\,\mathrm{sen}\,\theta\,\cos\,\psi - \dot{\theta}\,\mathrm{sen}\,\psi \ , \\ \omega_{z'} &= \dot{\psi} + \dot{\phi}\,\cos\,\theta \ . \end{aligned} \tag{3.4.30}$$

■ **Exercício 3.4.2.** Prove que

$$\begin{aligned} \omega_x &= \dot{\theta}\,\cos\,\phi + \dot{\psi}\,\mathrm{sen}\,\theta\,\mathrm{sen}\,\phi \ , \\ \omega_y &= \dot{\theta}\,\mathrm{sen}\,\phi - \dot{\psi}\,\mathrm{sen}\,\theta\,\cos\,\phi \ , \\ \omega_z &= \dot{\phi} + \dot{\psi}\,\cos\,\theta \end{aligned} \tag{3.4.31}$$

são as componentes da velocidade angular ao longo dos eixos fixos no espaço. ■

3.5 Grupo de Rotações e Geradores Infinitesimais

Com a ajuda da Eq.(3.4.11a) podemos escrever a Eq.(3.4.7) na forma

$$g'_\Sigma = \mathcal{A} g_\Sigma = (\boldsymbol{I} + \boldsymbol{J}_1\, d\Omega_1 + \boldsymbol{J}_2\, d\Omega_2 + \boldsymbol{J}_3\, d\Omega_3)\, g_\Sigma \ , \tag{3.5.1}$$

onde

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \ , \ J_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \ , \ J_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ . \qquad (3.5.2)$$

As matrizes J_i são chamadas de *geradores infinitesimais*[4] do grupo de rotações $SO(3)$.

■ **Exercício 3.5.1.** Prove que as matrizes J_i , $i = 1, 2, 3$, satisfazem a álgebra

$$\left[J_i, J_j\right] = J_k \ , \qquad (3.5.3a)$$

onde (i, j, k) é uma permutação cíclica de $(1, 2, 3)$ e $[\mathcal{A}, \mathcal{B}] = \mathcal{AB} - \mathcal{BA}$ é o comutador das matrizes $\mathcal{A}$ e $\mathcal{B}$. Mostre que, equivalentemente, pode-se escrever

$$\left[J_i, J_j\right] = \sum_{k=1}^{3} \epsilon_{ijk} J_k \ , \qquad (3.5.3b)$$

onde ϵ_{ijk} é o símbolo de Levi-Civita definido no Apêndice A. ■

A matriz que executa uma rotação finita de um vetor (sentido anti-horário e ponto de vista ativo) de um ângulo α em torno da direção $\hat{\mathbf{n}}$ pode ser obtida aplicando N rotações infinitesimais sucessivas do mesmo ângulo α/N e passando ao limite $N \to \infty$:

$$g'_\Sigma = \lim_{N\to\infty} \left(I + \frac{\alpha}{N} \hat{\mathbf{n}} \cdot \boldsymbol{J}\right)^N g_\Sigma = e^{\alpha \hat{\mathbf{n}} \cdot \boldsymbol{J}} g_\Sigma \ . \qquad (3.5.4)$$

Assim, demonstramos que uma rotação finita exprime-se em termos da função exponencial da matriz $\hat{\mathbf{n}} \cdot \boldsymbol{J} = n_1 J_1 + n_2 J_2 + n_3 J_3$. Nada há de misterioso na exponencial de uma matriz. Por definição,

$$e^{\mathcal{A}} = I + \mathcal{A} + \frac{\mathcal{A}^2}{2!} + \frac{\mathcal{A}^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{\mathcal{A}^k}{k!} \ , \qquad (3.5.5)$$

onde $\mathcal{A}$ e suas potências são matrizes $n \times n$. A série (3.5.5) é sempre convergente (Courant & Hilbert 1953).

A matriz J_k é o gerador das rotações infinitesimais em torno do k-ésimo eixo cartesiano e todas as matrizes do grupo de rotações tridimensionais $SO(3)$ podem ser representadas na forma $\mathcal{R}(\boldsymbol{\alpha}) = e^{\boldsymbol{\alpha} \cdot \boldsymbol{J}}$, onde $\boldsymbol{\alpha} \equiv \alpha \hat{\mathbf{n}}$. As matrizes $\mathcal{R}(\boldsymbol{\alpha})$ formam um *grupo de Lie* cujos geradores obedecem à *álgebra de Lie* associada ao grupo, Eqs.(3.5.3). As quantidades ϵ_{ijk} que aparecem na Eq.(3.5.3b) são chamadas de *constantes de estrutura* da álgebra de Lie do grupo de rotações tridimensionais.

[4] A rigor, as matrizes J_i constituem uma representação particular dos geradores do grupo $SO(3)$, a chamada representação fundamental.

■ **Exercício 3.5.2.** Usando a identidade

$$\det e^{\mathcal{A}} = e^{\mathrm{tr}\,\mathcal{A}} \ , \tag{3.5.6}$$

confirme que todas as matrizes do grupo $SO(3)$ têm determinante igual a 1. ■

3.6 Dinâmica em Referenciais Não-Inerciais

Sejam Σ e Σ' sistemas de referência *com origens coincidentes*, sendo Σ um sistema inercial e Σ' um sistema girante com velocidade angular $\boldsymbol{\omega}$. Seja $\mathbf{r}$ o vetor posição de uma partícula de massa m cuja equação de movimento em Σ escreve-se

$$m\left(\frac{d^2\mathbf{r}}{dt^2}\right)_{\Sigma} = \mathbf{F} \ , \tag{3.6.1}$$

onde $\mathbf{F}$ é a resultante das forças sobre a partícula. Usando a Eq.(3.4.14) temos

$$\mathbf{v}_{in} \equiv \left(\frac{d\mathbf{r}}{dt}\right)_{\Sigma} = \left(\frac{d\mathbf{r}}{dt}\right)_{\Sigma'} + \boldsymbol{\omega}\times\mathbf{r} \equiv \mathbf{v} + \boldsymbol{\omega}\times\mathbf{r} \ , \tag{3.6.2}$$

onde $\mathbf{v}_{in}$ é a velocidade da partícula no referencial inercial Σ e $\mathbf{v}$ denota a velocidade da partícula no sistema de referência girante, pois, devido à coincidência das origens, $\mathbf{r}$ também é o vetor posição da partícula relativamente a Σ'. De (3.6.2), com o uso reiterado de (3.4.14), deduz-se

$$\left(\frac{d^2\mathbf{r}}{dt^2}\right)_{\Sigma} = \left(\frac{d\mathbf{v}_{in}}{dt}\right)_{\Sigma} = \left(\frac{d\mathbf{v}_{in}}{dt}\right)_{\Sigma'} + \boldsymbol{\omega}\times\mathbf{v}_{in} = \left(\frac{d\mathbf{v}}{dt}\right)_{\Sigma'} + \frac{d\boldsymbol{\omega}}{dt}\times\mathbf{r} + 2\boldsymbol{\omega}\times\mathbf{v} + \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) \ . \tag{3.6.3}$$

A Eq.(3.4.14) com $\boldsymbol{g} = \boldsymbol{\omega}$ mostra que a aceleração angular é a mesma nos dois referenciais, daí a desnecessidade de subscrito em $d\boldsymbol{\omega}/dt$ nesta última equação. Denotando por $\mathbf{a} = (d\mathbf{v}/dt)_{\Sigma'}$ a aceleração da partícula no referencial girante e usando (3.6.1), obtém-se, finalmente, a equação de movimento da partícula no sistema de referência girante:

$$m\mathbf{a} = \mathbf{F} + 2m\mathbf{v}\times\boldsymbol{\omega} + m\boldsymbol{\omega}\times(\mathbf{r}\times\boldsymbol{\omega}) + m\mathbf{r}\times\frac{d\boldsymbol{\omega}}{dt} \ . \tag{3.6.4}$$

Para um observador no referencial não-inercial Σ' tudo se passa como se a partícula estivesse sujeita a uma força efetiva que é a soma da força usual $\mathbf{F}$ com: um primeiro termo conhecido como *força de Coriolis*; um segundo termo, quadrático em $\boldsymbol{\omega}$, chamado de *força centrífuga*; um terceiro termo, proporcional à aceleração angular do sistema de referência não-inercial, às vezes chamado de *força de Euler*.

■ **Exercício 3.6.1.** Se as origens de Σ e Σ' não coincidirem (Fig. 3.6), mostre que, em lugar de (3.6.4), teremos

$$m\mathbf{a} = \mathbf{F} - m\mathbf{a}_{o'} + 2m\mathbf{v}\times\boldsymbol{\omega} + m\boldsymbol{\omega}\times(\mathbf{r}\times\boldsymbol{\omega}) + m\mathbf{r}\times\frac{d\boldsymbol{\omega}}{dt} \ , \tag{3.6.5}$$

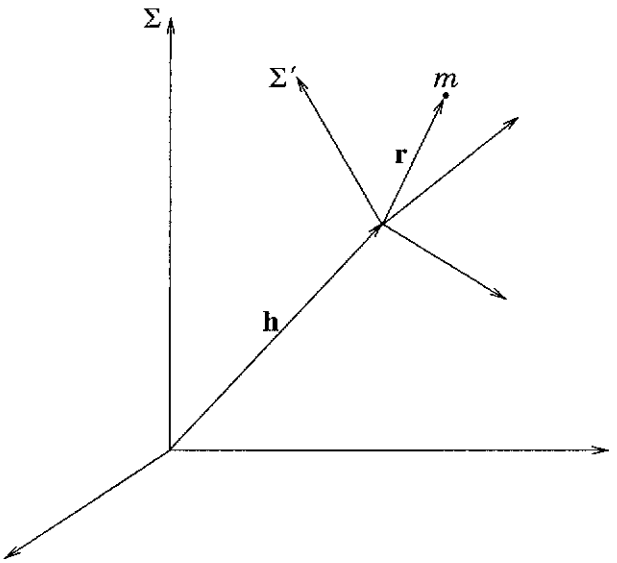


Figura 3.6: Rotação e translação simultâneas (Exercício 3.6.1).

onde $\mathbf{a}_{o'} = (d^2\mathbf{h}/dt^2)_\Sigma$ é a aceleração da origem de Σ' em relação a Σ. ■

Dinâmica no Referencial da Terra

Como uma das aplicações mais importantes dos desenvolvimentos anteriores, considere o movimento de uma partícula nas imediações da superfície da Terra. Será conveniente introduzir o sistema de eixos que acompanha a rotação da Terra indicado na Fig. 3.7. O eixo z é vertical, isto é, ortogonal à superfície da Terra; o eixo x é tangente ao meridiano apontando para o sul; o eixo y é tangente ao paralelo apontando para o leste. A latitude do local da superfície da Terra escolhido como origem do sistema cartesiano xyz é λ. Tomando a origem do sistema inercial Σ no centro da Terra, o vetor posição da origem do sistema girante é $\mathbf{R} = R\hat{\mathbf{k}}$, onde R é raio da Terra. Portanto, $\mathbf{R}$ é um vetor constante no referencial girante atado à superfície da Terra, e o uso repetido da Eq.(3.4.14) fornece imediatamente

$$\mathbf{a}_{o'} = m\left(\frac{d^2\mathbf{R}}{dt^2}\right)_\Sigma = \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{R}) \ , \tag{3.6.6}$$

onde usamos $d\boldsymbol{\omega}/dt = 0$, pois a velocidade angular da Terra é praticamente constante. Este resultado, substituído em (3.6.5), dá lugar a

$$m\mathbf{a} = \mathbf{T} + m\mathbf{g} + m\boldsymbol{\omega} \times (\mathbf{R} \times \boldsymbol{\omega}) + 2m\mathbf{v} \times \boldsymbol{\omega} + m\boldsymbol{\omega} \times (\mathbf{r} \times \boldsymbol{\omega}) \ , \tag{3.6.7}$$

onde tomamos $\mathbf{F} = m\mathbf{g} + \mathbf{T}$, com $\mathbf{T}$ denotando qualquer força adicional ao peso que esteja agindo sobre a partícula. A Eq.(3.6.7) pode ser escrita na forma

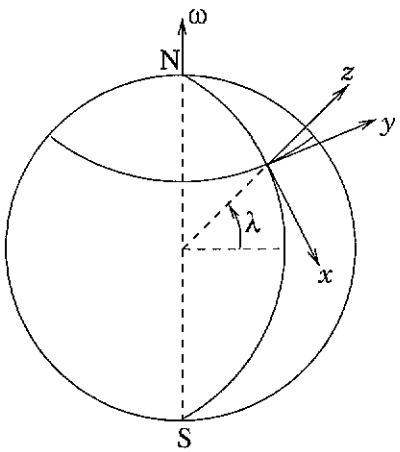


Figura 3.7: Referencial girante xyz fixo na superfície da Terra.

$$m\mathbf{a} = \mathbf{T} + m\mathbf{g}_{\text{ef}} + 2m\mathbf{v} \times \boldsymbol{\omega} + m\boldsymbol{\omega} \times (\mathbf{r} \times \boldsymbol{\omega}) \ , \tag{3.6.8}$$

onde

$$\mathbf{g}_{\text{ef}} = \mathbf{g} + \boldsymbol{\omega} \times (\mathbf{R} \times \boldsymbol{\omega}) \tag{3.6.9}$$

é a aceleração da gravidade efetiva. Para justificar esta denominação, considere um fio de prumo, isto é, uma partícula de massa m suspensa por um fio sujeito à tensão $\mathbf{T}$ e em equilíbrio na origem ($\mathbf{r} \equiv 0$) do sistema de coordenadas. Pela Eq.(3.6.8)

$$\mathbf{T} + m\mathbf{g}_{\text{ef}} = 0 \ , \tag{3.6.10}$$

ou seja, a direção do fio é determinada pelo vetor aceleração da gravidade efetiva. O vetor $\mathbf{g}_{\text{ef}}$ não aponta exatamente na direção do centro da Terra. No sistema de coordenadas da Fig. 3.7 temos $\mathbf{g} = -g\hat{\mathbf{k}}$ e $\boldsymbol{\omega} = -\omega \cos\lambda\,\hat{\mathbf{i}} + \omega \,\text{sen}\,\lambda\,\hat{\mathbf{k}}$, donde

$$\mathbf{g}_{\text{ef}} = -(g - R\omega^2 \cos^2\lambda)\,\hat{\mathbf{k}} + R\omega^2 \,\text{sen}\,\lambda\, \cos\lambda\,\hat{\mathbf{i}} \ . \tag{3.6.11}$$

Há, portanto, uma redução da aceleração da gravidade na direção vertical e o aparecimento de uma componente na direção norte-sul que faz com que $\mathbf{g}_{\text{ef}}$ não aponte exatamente na direção vertical, isto é, para o centro da Terra. Os efeitos, todavia, são pequenos porque

$$\omega \approx \frac{2\pi.}{24 \times 3600\,\text{s}} \approx 7,3 \times 10^{-5}\,\text{s}^{-1}$$

e, conseqüentemente,

$$R\omega^2 \approx (6400\,\mathrm{km}) \times (7,3 \times 10^{-5}\,\mathrm{s}^{-1})^2 \approx 3 \times 10^{-2}\,\mathrm{m/s}^2 \ ,$$

que é cerca de 300 vezes menor do que a aceleração da gravidade. Este efeito centrífugo, embora diminuto, não pode ser desprezado nos estudos da forma da Terra ou do movimento de foguetes.

Para objetos em movimento nas proximidades da superfície da Terra, o último termo no segundo membro da Eq.(3.6.7) é usualmente muito menor do que o termo correspondente à força de Coriolis. Essencialmente, o efeito centrífugo resume-se em converter g em g_{ef}, e a principal alteração no movimento induzida pela rotação da Terra provém da força de Coriolis. Para um avião supersônico com velocidade $v \approx 2000$km/h, a aceleração de Coriolis não é superior a $2\omega v \approx 0,08\mathrm{m/s}^2$, ou cerca de $0,008g$. Embora pequena, a aceleração de Coriolis é importante em muitas circunstâncias, como veremos em alguns exemplos que se seguirão. No hemisfério norte a força de Coriolis tende a defletir para a direita um objeto disparado paralelamente à superfície da Terra (horizontalmente), a deflexão sendo para a esquerda no hemisfério sul. Presume-se que este efeito explique o comportamento dos ciclones, que tipicamente giram no sentido anti-horário no hemisfério norte mas, no hemisfério sul, giram no sentido horário (Goldstein 1980).

■ **Exemplo 3.6.1.** Um objeto cai de uma altura h a partir do repouso. Determinar a deflexão transversal do objeto ao atingir o solo, provocada pela força de Coriolis.

Solução. Se o objeto cai verticalmente, $\mathbf{v} = -v\hat{\mathbf{k}}$ com $v > 0$ e a força de Coriolis é

$$\mathbf{F}_{\mathrm{Cor}} = 2m\mathbf{v} \times \boldsymbol{\omega} = -2mv\hat{\mathbf{k}} \times (-\omega \cos\lambda\,\hat{\mathbf{i}} + \omega\,\mathrm{sen}\,\lambda\,\hat{\mathbf{k}}) = 2m\omega v \cos\lambda\,\hat{\mathbf{j}} \ . \tag{3.6.12}$$

O desvio, portanto, é no sentido oeste-leste. No hemisfério sul $\lambda < 0$ mas o sentido da deflexão é o mesmo, pois $\cos(-\lambda) = \cos\lambda$. Na direção vertical a equação de movimento é aproximadamente

$$\ddot{z} = -g \implies \dot{z} = -gt \implies z = h - \frac{1}{2}gt^2 \ , \tag{3.6.13}$$

com as condições iniciais $\dot{z}(0) = 0, z(0) = h$. Na direção oeste-leste temos

$$\ddot{y} = 2\omega v \cos\lambda \approx 2\omega g t \cos\lambda \ . \tag{3.6.14}$$

A rigor, só imediatamente após o início da queda o vetor velocidade $\mathbf{v}$ é vertical. À medida que cai, a partícula adquire uma velocidade transversal que permanece sempre muito menor do que a componente vertical porque a aceleração de Coriolis é muito menor do que a aceleração da gravidade. Isto justifica aproximar v apenas por sua componente vertical durante todo o movimento de queda livre, como fizemos na Eq.(3.6.14). Com as condições iniciais $y(0) = \dot{y}(0) = 0$ resulta

$$y = \frac{\omega g}{3}t^3 \cos\lambda = \frac{\omega}{3}\sqrt{\frac{8(h-z)^3}{g}}\cos\lambda \ . \tag{3.6.15}$$

Ao atingir o solo ($z = 0$), o deslocamento transversal é

$$\Delta y = \frac{\omega}{3}\sqrt{\frac{8h^3}{g}}\cos\lambda \ . \tag{3.6.16}$$

No equador, onde o efeito é máximo, para uma queda de uma altura $h = 100$m a deflexão é

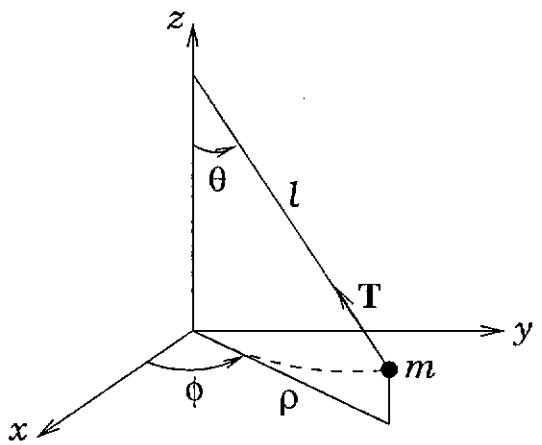


Figura 3.8: Pêndulo de Foucault.

$\Delta y \approx 2,2$cm. Apesar de o desvio ser mensurável, a experiência é de difícil realização, pois é preciso eliminar influências de ventos, resistência do ar, e outras que podem mascarar a deflexão de Coriolis. ■

■ **Exemplo 3.6.2.** Determinar o efeito da rotação da Terra sobre um pêndulo cujo ponto de suspensão pode girar livremente, conhecido como pêndulo de Foucault.

Solução. Desprezando o termo quadrático em ω em (3.6.7) e tomando $\mathbf{g}_{\text{ef}} \approx \mathbf{g}$, a equação de movimento de um pêndulo simples torna-se, aproximadamente,

$$m\mathbf{a} = \mathbf{T} + m\mathbf{g} + 2m\mathbf{v} \times \boldsymbol{\omega} \ , \tag{3.6.17}$$

onde $\mathbf{T}$ denota a tensão no fio. É conveniente escolher a origem do sistema de coordenadas no ponto de equilíbrio do pêndulo (Fig. 3.8). No caso de pequenas oscilações, o deslocamento vertical do pêndulo é muito menor do que o deslocamento horizontal. De fato,

$$z = l(1 - \cos\theta) \approx l\frac{\theta^2}{2} \ ,$$

de modo que, com $\rho = l \operatorname{sen}\theta \approx l\theta$,

$$\frac{z}{\rho} \approx \frac{\theta}{2} \ll 1$$

para θ pequeno. Portanto, na direção vertical o pêndulo permanece praticamente em repouso, a tensão no fio equilibrando o peso:

$$T\cos\theta \approx mg \implies T \approx mg \tag{3.6.18}$$

pois $\cos\theta \approx 1$. A força de Coriolis é transversal ao plano de oscilação do pêndulo, o que faz com que o pêndulo não oscile num plano fixo. As componentes horizontais de (3.6.17) são

$$\ddot{x} = -\frac{T}{m}\operatorname{sen}\theta\cos\phi + 2\omega_z\dot{y} \ , \quad \ddot{y} = -\frac{T}{m}\operatorname{sen}\theta\operatorname{sen}\phi - 2\omega_z\dot{x} \ , \tag{3.6.19}$$

com

$$\omega_z = \omega \operatorname{sen} \lambda \ . \tag{3.6.20}$$

Mas

$$\operatorname{sen}\theta = \frac{\rho}{l} \quad , \quad \rho\cos\phi = x \quad , \quad \rho\operatorname{sen}\phi = y \ , \tag{3.6.21}$$

de modo que, com o uso de (3.6.18), (3.6.19) assume a forma

$$\ddot{x} + \omega_0^2 x = 2\omega_z \dot{y} \quad , \quad \ddot{y} + \omega_0^2 y = -2\omega_z \dot{x} \ , \tag{3.6.22}$$

onde

$$\omega_0 = \sqrt{\frac{g}{l}} \ . \tag{3.6.23}$$

Note que para um pêndulo típico com $l < 10\text{m}$ temos $\omega_0 > 1\,\text{s}^{-1} \gg \omega_z$. As equações de movimento do pêndulo resolvem-se mais facilmente com a ajuda da variável complexa $\zeta(t) = x(t) + iy(t)$, pois as equações (3.6.22) são as partes real e imaginária de

$$\ddot{\zeta} + 2i\omega_z \dot{\zeta} + \omega_0^2 \zeta = 0 \ . \tag{3.6.24}$$

Esta última equação admite soluções da forma $\zeta(t) = \zeta_0\, e^{-ipt}$ com

$$-p^2 + 2\omega_z p + \omega_0^2 = 0 \implies p = \omega_z \pm \sqrt{\omega_z^2 + \omega_0^2} \approx \omega_z \pm \omega_0 \ . \tag{3.6.25}$$

A solução geral para $\zeta(t)$ é

$$\zeta(t) = Ae^{-i(\omega_z - \omega_0)t} + Be^{-i(\omega_z + \omega_0)t} \ . \tag{3.6.26}$$

Adotando as condições iniciais $x(0) = a$, $\dot{x}(0) = 0$, $y(0) = \dot{y}(0) = 0$, de modo que o pêndulo começa oscilando no plano xz, obtém-se

$$A + B = a \quad , \quad B = \frac{\omega_0 - \omega_z}{\omega_0 + \omega_z} A \approx \frac{\omega_0}{\omega_0} A = A \implies A = B = a/2 \ , \tag{3.6.27}$$

donde

$$\zeta(t) = a\cos\omega_0 t\, e^{-i\omega_z t} \ . \tag{3.6.28}$$

Portanto,

$$x(t) = \operatorname{Re}\zeta(t) = a\cos\omega_z t\cos\omega_0 t \quad , \quad y(t) = \operatorname{Im}\zeta(t) = -a\operatorname{sen}\omega_z t\cos\omega_0 t \ . \tag{3.6.29}$$

O ângulo φ que o plano de oscilação faz com o plano inicial de oscilação (plano xz) satisfaz

$$\tan\varphi = \frac{y}{x} = -\tan\omega_z t \implies \varphi = -\omega_z t = -(\omega\operatorname{sen}\lambda)t \ . \tag{3.6.30}$$

O plano de oscilação gira com velocidade angular $\omega \operatorname{sen} \lambda$ em torno da vertical,[5] o sentido, visto de cima, sendo horário no hemisfério norte ($\lambda > 0$) e anti-horário no hemisfério sul ($\lambda < 0$). A cada 24 horas o plano de oscilação gira de $2\pi \operatorname{sen} \lambda$. Na latitude do Rio de Janeiro ($\lambda = -22°54'$), o plano de oscilação do pêndulo gira de 140^o por dia no sentido anti-horário. No pólo norte a rotação é de 360^o por dia e pode ser entendida da seguinte maneira: o plano de oscilação é fixo num referencial inercial, mas a Terra gira de 360^o por dia no sentido anti-horário embaixo do pêndulo, de modo que, para um observador na Terra, o sentido de rotação do plano de oscilação é horário. Como observação final, note que (3.6.29) mostra que a ordem de grandeza de $\dot{x}$ e $\dot{y}$ é $\omega_0 a$ porque $\omega_0 \gg \omega > \omega_z$. Por outro lado, a ordem de grandeza de $|\boldsymbol{\omega} \times (\mathbf{r} \times \boldsymbol{\omega})|$ é $\omega^2 a$. Portanto, a razão entre o termo centrífugo e o termo de Coriolis em (3.6.7) é $\omega^2 a / 2\omega_0 a\omega = \omega / 2\omega_o \ll 1$, o que justifica *a posteriori* a aproximação feita ao se escrever (3.6.17). ■

[5] Foucault intuiu este resultado em 1851, antes da análise matemática do problema levando em conta a força de Coriolis como causa da rotação (Dugas 1988).

PROBLEMAS

3.1. Em três dimensões a Eq.(3.4.11b) estabelece uma correspondência biunívoca entre vetores e matrizes anti-simétricas reais. (i) Se $\mathcal{A} = (a_{kl})$ é a matriz associada ao vetor $\mathbf{v} = (v_1, v_2, v_3)$, mostre que $a_{kl} = \sum_{m=1}^{3} \epsilon_{kml} v_m$, onde ϵ_{kml} é o símbolo de Levi-Civita definido no Apêndice A. Mostre que, inversamente, $v_m = \frac{1}{2} \sum_{k,l=1}^{3} \epsilon_{kml} a_{kl}$. (ii) Prove que os autovalores da matriz anti-simétrica associada ao vetor $\mathbf{n}$ são zero e $\pm i|\mathbf{n}|$. (iii) Se $\mathcal{A}$ é uma matriz real anti-simétrica, prove que as matrizes $\boldsymbol{I} \pm \mathcal{A}$ são não-singulares e que a matriz $\mathcal{B} = (\boldsymbol{I} + \mathcal{A})(\boldsymbol{I} - \mathcal{A})^{-1}$ é ortogonal.

3.2. Exprima o vetor $\mathbf{r}'$ obtido por reflexão do vetor $\mathbf{r}$ em relação a um plano cujo vetor unitário normal é $\hat{\mathbf{n}}$. Sem efetuar nenhum cálculo, usando apenas argumentos geométricos, determine os autovalores e autovetores da matriz de transformação correspondente $\mathcal{A}$. Se $\hat{\mathbf{n}} = (n_1, n_2, n_3)$, mostre que $\mathcal{A}$ tem elementos $A_{ij} = \delta_{ij} - 2n_i\, n_j$ e é uma matriz ortogonal imprópria.

3.3. Para a matriz $\mathcal{D}$ da Eq.(3.3.1), prove que

$$\mathcal{D}^n = \begin{pmatrix} \cos n\phi & \operatorname{sen} n\phi & 0 \\ -\operatorname{sen} n\phi & \cos n\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

3.4. Qualquer matriz ortogonal própria $\mathcal{A}$ corresponde a uma rotação de um certo ângulo Φ em torno de uma certa direção. O eixo de rotação é determinado pelo autovetor de $\mathcal{A}$ com autovalor 1. Se escolhermos as coordenadas de tal modo que o eixo z coincida com o referido autovetor, a matriz $\mathcal{A}$ terá a forma (3.3.1). Considerando que o traço de uma matriz é independente do sistema de coordenadas, prove que o ângulo de rotação é dado por $\cos\Phi = (\operatorname{tr}\mathcal{A} - 1)/2$.

3.5. Uma esfera de raio R rola sem deslizar sobre uma superfície plana. Se (x, y, R) são as coordenadas do centro da esfera, mostre que, em termos dos ângulos de Euler, as condições de rolamento sem deslizamento são

$$\dot{x} - R(\dot{\theta}\operatorname{sen}\phi - \dot{\psi}\operatorname{sen}\theta\cos\phi) = 0 \quad , \quad \dot{y} + R(\dot{\theta}\cos\phi + \dot{\psi}\operatorname{sen}\theta\operatorname{sen}\phi) = 0$$

e prove que esses vínculos não são holônomos.

3.6. O deslocamento linear de um corpo pode ser representado por um vetor, e a velocidade linear é a derivada do vetor posição em relação ao tempo. O vetor velocidade angular, no entanto, não é, em geral, a derivada temporal de um vetor deslocamento angular. Para provar isso, suponha que exista um vetor $\boldsymbol{\Omega}$ cujas componentes $\Omega_x, \Omega_y, \Omega_z$ sejam funções dos ângulos de Euler tais que as respectivas derivadas temporais coincidam com as componentes correspondentes da velocidade angular, isto é,

$$\omega_x = \frac{\partial\Omega_x}{\partial\theta}\dot{\theta}+\frac{\partial\Omega_x}{\partial\phi}\dot{\phi}+\frac{\partial\Omega_x}{\partial\psi}\dot{\psi}\ ,\ \omega_y = \frac{\partial\Omega_y}{\partial\theta}\dot{\theta}+\frac{\partial\Omega_y}{\partial\phi}\dot{\phi}+\frac{\partial\Omega_y}{\partial\psi}\dot{\psi}\ ,\ \omega_z = \frac{\partial\Omega_z}{\partial\theta}\dot{\theta}+\frac{\partial\Omega_z}{\partial\phi}\dot{\phi}+\frac{\partial\Omega_z}{\partial\psi}\dot{\psi}\ .$$

Usando (3.4.31), prove que não existe nenhum vetor $(\Omega_x, \Omega_y, \Omega_z)$ que satisfaça as equações acima.

3.7. A Eq.(3.4.14) pode ser deduzida de forma inteiramente analítica, sem apelar para a definição geométrica do vetor velocidade angular. (i) Por definição, a derivada de uma matriz é a matriz constituída pelas derivadas de seus elementos. Se $\boldsymbol{\mathcal{A}}$ é uma matriz ortogonal dependente do tempo, prove, diferenciando a identidade $\boldsymbol{\mathcal{A}}^T\boldsymbol{\mathcal{A}} = \boldsymbol{I}$ em relação ao tempo, que a matriz $\dot{\boldsymbol{\mathcal{A}}}^T\boldsymbol{\mathcal{A}}$ é anti-simétrica, isto é, $\dot{\boldsymbol{\mathcal{A}}}^T\boldsymbol{\mathcal{A}} = -\left(\dot{\boldsymbol{\mathcal{A}}}^T\boldsymbol{\mathcal{A}}\right)^T$. (ii) Sejam $\mathbf{r}_\Sigma$ e $\mathbf{r}_{\Sigma'}$ as componentes do *mesmo vetor* posição $\mathbf{r}$ de uma partícula relativamente aos referenciais inercial Σ e girante Σ', e $\boldsymbol{\mathcal{A}}$ a matriz de rotação de Σ para Σ'. Se $\mathbf{v}$ e $\mathbf{v}'$ são as velocidades da partícula em relação aos referenciais inercial e girante, respectivamente, temos $\mathbf{v}_\Sigma = d\mathbf{r}_\Sigma/dt$ e $\mathbf{v}'_{\Sigma'} = d\mathbf{r}_{\Sigma'}/dt$. Partindo de $\mathbf{r}_\Sigma = \boldsymbol{\mathcal{A}}^T\mathbf{r}_{\Sigma'}$, deduza

$$\mathbf{v}_\Sigma = \mathbf{v}'_\Sigma + \dot{\boldsymbol{\mathcal{A}}}^T\boldsymbol{\mathcal{A}}\,\mathbf{r}_\Sigma\ .$$

Levando em conta o item (i), a correspondência (3.4.11b) entre vetores e matrizes anti-simétricas e que uma igualdade entre vetores é válida em qualquer sistema de coordenadas, prove que existe um vetor $\boldsymbol{\omega}$ tal que

$$\mathbf{v} = \mathbf{v}' + \boldsymbol{\omega}\times\mathbf{r}\ ,$$

que coincide com a Eq.(3.4.14).

3.8. Seja $\boldsymbol{\mathcal{P}}$ uma matriz de rotação de 180^o em torno de um eixo arbitrário. (ii) Determine $\boldsymbol{\mathcal{P}}^2$ sem cálculos, refletindo sobre o seu significado. (ii) Sendo $\boldsymbol{\mathcal{A}} = (\boldsymbol{I}+\boldsymbol{\mathcal{P}})/2$ e $\boldsymbol{\mathcal{B}} = (\boldsymbol{I}-\boldsymbol{\mathcal{P}})/2$, prove que $\boldsymbol{\mathcal{A}}^2 = \boldsymbol{\mathcal{A}}$ e $\boldsymbol{\mathcal{B}}^2 = \boldsymbol{\mathcal{B}}$. (iii) Mostre que as matrizes $\boldsymbol{\mathcal{A}}$ e $\boldsymbol{\mathcal{B}}$ são singulares e calcule o seu produto.

3.9. Uma partícula é disparada verticalmente com velocidade inicial v_0, atinge uma altura máxima e retorna ao solo. Mostre que a deflexão provocada pela força de Coriolis, quando ela atinge novamente o chão, tem sentido oposto e é quatro vezes maior do que o desvio produzido quando a partícula é largada em repouso da mesma altura máxima.

3.10. Mostre que a equação de movimento (3.6.5) pode ser obtida da lagrangiana

$$L = \frac{mv^2}{2} - V - m\mathbf{r}\cdot\mathbf{a}_{o'} + \frac{m}{2}(\boldsymbol{\omega}\times\mathbf{r})^2 + m\mathbf{v}\cdot(\boldsymbol{\omega}\times\mathbf{r})\ ,$$

onde $\mathbf{v} = \dot{\mathbf{r}}$ e V é o potencial da força $\mathbf{F}$. Obtenha a integral de Jacobi associada a esta lagrangiana no caso em que $\mathbf{a}_{o'}$ e $\boldsymbol{\omega}$ são constantes.

Capítulo 4

DINÂMICA DO CORPO RÍGIDO

Roda mundo, roda-gigante
Roda moinho, roda pião
O tempo rodou num instante
Nas voltas do meu coração.
Chico Buarque, *Roda Viva*

Agora que estamos de posse dos instrumentos cinemáticos necessários, podemos proceder ao exame de alguns problemas simples mas importantes da dinâmica do corpo rígido, o que exige levar em conta as causas do movimento: forças e torques. Duas grandezas físicas essenciais ao estudo do movimento de um corpo rígido são o momento angular e a energia cinética, que passamos a considerar em detalhe.

4.1 Momento Angular e Tensor de Inércia

As equações de movimento de um corpo rígido podem ser escritas na forma

$$\frac{d\mathbf{P}}{dt} = \mathbf{F} \ , \tag{4.1.1}$$

$$\frac{d\mathbf{L}}{dt} = \mathbf{N} \ , \tag{4.1.2}$$

onde $\mathbf{F}$ é a força resultante e $\mathbf{N}$ é o torque total, ao passo que $\mathbf{P}$ e $\mathbf{L}$ são o momento linear total e o momento angular total, respectivamente. Como vimos na Seção 1.1, as equações (4.1.1) e (4.1.2) são verdadeiras somente se as taxas de variação temporal são relativas a um referencial inercial e se são obedecidas certas restrições quanto ao ponto em relação ao qual o torque e o momento angular são calculados. Em particular, (4.1.2) vigora se o ponto de referência para o cálculo de $\mathbf{N}$ e $\mathbf{L}$ é um ponto imóvel (conveniente no caso em que o corpo rígido está limitado a girar em torno de um ponto fixo) ou é o centro de massa do corpo.

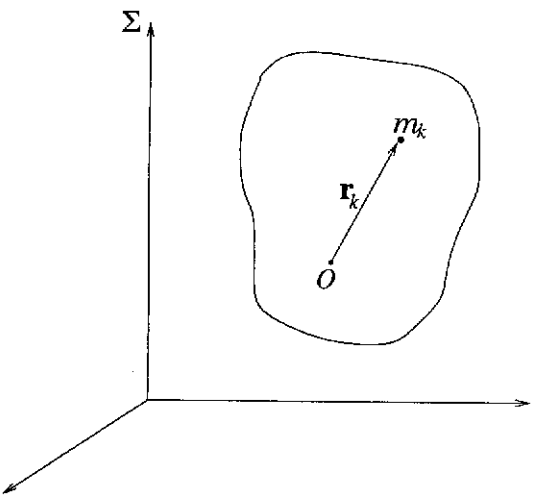


Figura 4.1: O ponto O é um ponto do corpo que permanece fixo relativamente ao referencial inercial Σ ou é o centro de massa do corpo.

Seja O um ponto fixo ou o centro de massa de um corpo rígido. Do ponto de vista do sistema de referência inercial Σ (Fig. 4.1), o momento angular total do corpo em relação ao ponto O é

$$\mathbf{L} = \sum_{k=1}^{N} \mathbf{r}_k \times \mathbf{p}_k = \sum_{k=1}^{N} m_k \, \mathbf{r}_k \times \mathbf{v}_k \ , \tag{4.1.3}$$

onde N é o número de partículas do corpo e $\mathbf{v}_k = (d\mathbf{r}_k/dt)_\Sigma$ é a velocidade da k-ésima partícula em relação ao ponto O, do ponto de vista do sistema inercial Σ . Mas $\mathbf{r}_k$ é um vetor constante relativamente a um sistema cartesiano Σ' fixo no corpo, donde

$$\mathbf{v}_k = \left(\frac{d\mathbf{r}_k}{dt}\right)_\Sigma = \left(\frac{d\mathbf{r}_k}{dt}\right)_{\Sigma'} + \ \boldsymbol{\omega} \times \mathbf{r}_k \ = \ \boldsymbol{\omega} \times \mathbf{r}_k \ . \tag{4.1.4}$$

Com este último resultado nos é permitido escrever

$$\mathbf{L} = \sum_{k=1}^{N} m_k \, \mathbf{r}_k \times (\boldsymbol{\omega} \times \mathbf{r}_k) = \sum_{k=1}^{N} m_k \left[\, r_k^2 \, \boldsymbol{\omega} - (\mathbf{r}_k \cdot \boldsymbol{\omega}) \mathbf{r}_k \, \right] \ , \tag{4.1.5}$$

onde usamos a identidade $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c}$. Neste estágio torna-se conveniente adotar uma notação indicial para as componentes de $\mathbf{r}_k$, isto é, $\mathbf{r}_k = (x_1^{(k)}, x_2^{(k)}, x_3^{(k)})$, que corresponde, na notação tradicional, a $x_1^{(k)} \equiv x_k$, $x_2^{(k)} \equiv y_k$, $x_3^{(k)} \equiv z_k$. Escrevendo $\mathbf{L} = (L_1, L_2, L_3)$, a i-ésima componente de (4.1.5) é dada por

$$L_i = \Big(\sum_{k=1}^{N} m_k \, r_k^2\Big) \, \omega_i - \sum_{k=1}^{N} \sum_{j=1}^{3} m_k \, x_j^{(k)} \, \omega_j \, x_i^{(k)} = \sum_{k=1}^{N} \sum_{j=1}^{3} m_k \left[\, r_k^2 \, \delta_{ij} - x_i^{(k)} x_j^{(k)} \, \right] \omega_j \ , \tag{4.1.6}$$

já que

$$\mathbf{r}_k \cdot \boldsymbol{\omega} = \sum_{j=1}^{3} x_j^{(k)}\, \omega_j \quad , \qquad \omega_i = \sum_{j=1}^{3} \delta_{ij}\, \omega_j \ . \tag{4.1.7}$$

Conseqüentemente,

$$L_i = \sum_{j=1}^{3} I_{ij}\, \omega_j \ , \tag{4.1.7}$$

onde

$$I_{ij} = \sum_{k=1}^{N} m_k \left[r_k^2\, \delta_{ij} - x_i^{(k)}\, x_j^{(k)} \right] \ . \tag{4.1.8}$$

As nove quantidades I_{ij}, $i, j = 1, 2, 3$, são chamadas de componentes do *tensor de inércia.*

Na maior parte dos problemas envolvendo corpos rígidos é mais prático tratá-los como corpos contínuos do que como sistemas discretos de partículas. A transição da representação discreta para a contínua faz-se mediante a correspondência

$$\mathbf{r}_k \quad \longrightarrow \quad \mathbf{r} \qquad , \qquad \sum_{k=1}^{N} m_k \quad \longrightarrow \quad \int_V \rho dv \ . \tag{4.1.9}$$

Em outras palavras, a massa puntiforme m_k na posição $\mathbf{r}_k$ é substituída pelo elemento de massa $dm = \rho dv$ localizado no ponto $\mathbf{r}$, e a soma sobre todas as partículas é substituída por uma integração estendida ao volume V ocupado pelo corpo. A versão contínua de (4.1.8) é, portanto,

$$I_{ij} = \int_V \left[r^2\, \delta_{ij} - x_i\, x_j \right] \rho\, dv \ . \tag{4.1.10}$$

Uma representação intrínseca do tensor de inércia que remove qualquer referência a um sistema de coordenadas particular é obtida notando que a equação (4.1.5) pode ser escrita na forma simbólica

$$\mathbf{L} = \Big(\sum_{k=1}^{N} m_k\, r_k^2\Big)\, \boldsymbol{\omega} - \Big(\sum_{k=1}^{N} m_k\, \mathbf{r}_k\, \mathbf{r}_k\Big) \cdot \boldsymbol{\omega} \ , \tag{4.1.11}$$

onde a segunda soma envolve um *produto tensorial* ou *produto diádico* de vetores, cujo significado e propriedades passamos a discutir.

4.2 Interlúdio Matemático: Tensores e Diádicos

O produto diádico ou produto tensorial de dois vetores **A** e **B**, denotado simplesmente[1] por **AB**, é um operador linear que transforma vetores em vetores. O operador

[1] O produto diádico é representado pela mera justaposição de dois vetores, sem indicação de produto escalar nem de produto vetorial. Outra notação freqüentemente empregada é $\mathbf{A} \otimes \mathbf{B}$.

AB é também chamado de uma *díada*, e o resultado de sua ação sobre um vetor qualquer **C** define-se por

$$(\mathbf{AB}) \cdot \mathbf{C} = \mathbf{A}(\mathbf{B} \cdot \mathbf{C}) \ . \tag{4.2.1}$$

Como $\mathbf{B} \cdot \mathbf{C}$ é um escalar, o segundo membro de (4.2.1) é o produto de um escalar por um vetor, ou seja, é de fato um vetor. A linearidade é uma decorrência imediata desta definição :

$$\text{(i) } (\mathbf{AB}) \cdot (\alpha\mathbf{C} + \beta\mathbf{D}) = \alpha(\mathbf{AB}) \cdot \mathbf{C} + \beta(\mathbf{AB}) \cdot \mathbf{D} \ , \quad \alpha, \beta \in I\!R. \tag{4.2.2}$$

Por definição valem as seguintes propriedades adicionais (α e β são escalares):

$$\begin{aligned} &\text{(ii) } (\alpha\mathbf{AB} + \beta\mathbf{CD}) \cdot \mathbf{E} = \alpha(\mathbf{AB}) \cdot \mathbf{E} + \beta(\mathbf{CD}) \cdot \mathbf{E} \ ; \\ &\text{(iii) } (\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC} \ ; \\ &\text{(iv) } \mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC} \ . \end{aligned} \tag{4.2.3}$$

O produto tensorial é distributivo em relação à adição vetorial, mas não é comutativo, isto é, $\mathbf{AB} \neq \mathbf{BA}$ em geral. Também é possível definir o produto escalar de um vetor por uma díada pela esquerda:

$$\mathbf{C} \cdot (\mathbf{AB}) = (\mathbf{C} \cdot \mathbf{A})\mathbf{B} \ . \tag{4.2.4}$$

Por este motivo, uma notação bastante sugestiva para uma díada é uma letra encimada por uma seta dupla, assim: $\overleftrightarrow{\mathbf{T}} = \mathbf{AB}$.

O *tensor*[2] ou o *diádico* mais geral possível é uma soma de díadas:

$$\overleftrightarrow{\mathbf{T}} = \mathbf{AB} + \mathbf{CD} + \cdots \ . \tag{4.2.5}$$

Decompondo os vetores $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \ldots$ em termos dos vetores unitários de uma base ortonormal $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$ verificamos que o tensor mais geral $\overleftrightarrow{\mathbf{T}}$ escreve-se

$$\begin{aligned} \overleftrightarrow{\mathbf{T}} = {} & T_{11}\,\hat{\mathbf{e}}_1\hat{\mathbf{e}}_1 + T_{12}\,\hat{\mathbf{e}}_1\hat{\mathbf{e}}_2 + T_{13}\,\hat{\mathbf{e}}_1\hat{\mathbf{e}}_3 + \\ & T_{21}\,\hat{\mathbf{e}}_2\hat{\mathbf{e}}_1 + T_{22}\,\hat{\mathbf{e}}_2\hat{\mathbf{e}}_2 + T_{23}\,\hat{\mathbf{e}}_2\hat{\mathbf{e}}_3 + \\ & T_{31}\,\hat{\mathbf{e}}_3\hat{\mathbf{e}}_1 + T_{32}\,\hat{\mathbf{e}}_3\hat{\mathbf{e}}_2 + T_{33}\,\hat{\mathbf{e}}_3\hat{\mathbf{e}}_3 \ , \end{aligned} \tag{4.2.6}$$

onde $T_{11}, \ldots, T_{33}$ são nove números reais, chamados de componentes do tensor $\overleftrightarrow{\mathbf{T}}$ na base em questão. Claramente,

$$T_{ij} = \hat{\mathbf{e}}_i \cdot \overleftrightarrow{\mathbf{T}} \cdot \hat{\mathbf{e}}_j \ , \tag{4.2.7}$$

[2]Tendo em vista que neste capítulo não necessitaremos de tensores de ordem superior, sempre que não houver perigo de confusão usaremos simplesmente "tensor" para nos referirmos a um tensor de segunda ordem (vide o antepenúltimo parágrafo desta seção) .

não havendo necessidade de parênteses nesta última equação porque $\mathbf{A}\cdot(\overleftrightarrow{\mathbf{T}}\cdot\mathbf{B}) = (\mathbf{A}\cdot\overleftrightarrow{\mathbf{T}})\cdot\mathbf{B}$ (verifique!).

O *tensor unidade* ou *identidade*, simbolizado por $\overleftrightarrow{\mathbf{1}}$, é definido por

$$\overleftrightarrow{\mathbf{1}} = \hat{\mathbf{e}}_1\hat{\mathbf{e}}_1 + \hat{\mathbf{e}}_2\hat{\mathbf{e}}_2 + \hat{\mathbf{e}}_3\hat{\mathbf{e}}_3 \ . \tag{4.2.8}$$

Dado um vetor arbitrário $\mathbf{A} = A_1\hat{\mathbf{e}}_1 + A_2\hat{\mathbf{e}}_2 + A_3\hat{\mathbf{e}}_3$ temos

$$\overleftrightarrow{\mathbf{1}}\cdot\mathbf{A} = \hat{\mathbf{e}}_1(\hat{\mathbf{e}}_1\cdot\mathbf{A}) + \hat{\mathbf{e}}_2(\hat{\mathbf{e}}_2\cdot\mathbf{A}) + \hat{\mathbf{e}}_3(\hat{\mathbf{e}}_3\cdot\mathbf{A}) = A_1\hat{\mathbf{e}}_1 + A_2\hat{\mathbf{e}}_2 + A_3\hat{\mathbf{e}}_3 = \mathbf{A}$$

com idêntico resultado para $\mathbf{A}\cdot\overleftrightarrow{\mathbf{1}}$, o que justifica o nome dado a $\overleftrightarrow{\mathbf{1}}$. Evidentemente, as componentes de $\overleftrightarrow{\mathbf{1}}$ na base ortonormal $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$ são

$$\hat{\mathbf{e}}_i\cdot\overleftrightarrow{\mathbf{1}}\cdot\hat{\mathbf{e}}_j = \hat{\mathbf{e}}_i\cdot(\overleftrightarrow{\mathbf{1}}\cdot\hat{\mathbf{e}}_j) = \hat{\mathbf{e}}_i\cdot\hat{\mathbf{e}}_j = \delta_{ij} \ ,$$

como era de se esperar. É fácil demonstrar que, projetada em componentes na base $\{\hat{\mathbf{e}}_i\}_{i=1}^3$, a equação

$$\mathbf{A} = \overleftrightarrow{\mathbf{T}}\cdot\mathbf{B} \tag{4.2.9a}$$

equivale a

$$A_i = \sum_{j=1}^{3} T_{ij}B_j \ . \tag{4.2.9b}$$

Com efeito, devido à linearidade,

$$\mathbf{A} = \overleftrightarrow{\mathbf{T}}\cdot\Big(\sum_{j=1}^{3}\hat{\mathbf{e}}_jB_j\Big) = \sum_{j=1}^{3}\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{e}}_j\,B_j \ ,$$

donde

$$A_i = \hat{\mathbf{e}}_i\cdot\mathbf{A} = \sum_{j=1}^{3}\hat{\mathbf{e}}_i\cdot\overleftrightarrow{\mathbf{T}}\cdot\hat{\mathbf{e}}_j\,B_j \ ,$$

que corresponde à Eq.(4.2.9b).

Operador de Rotação Finita

A notação diádica é particularmente adequada para obter uma representação intrínseca do operador linear que executa a rotação finita de um vetor em torno de um eixo

dado. Antes, porém, de fazê-lo, necessitamos de uma definição simples. O produto vetorial pela esquerda de um vetor $\mathbf{A}$ por um tensor $\overleftrightarrow{\mathbf{T}}$ é um outro tensor $\mathbf{A} \times \overleftrightarrow{\mathbf{T}}$ definido de modo natural por sua ação sobre um vetor arbitrário $\mathbf{B}$:

$$(\mathbf{A} \times \overleftrightarrow{\mathbf{T}}) \cdot \mathbf{B} = \mathbf{A} \times (\overleftrightarrow{\mathbf{T}} \cdot \mathbf{B}) \ . \tag{4.2.10}$$

■ **Exemplo 4.2.1.** Com o emprego de diádicos, exprimir o operador de rotação numa forma intrínseca, independente de qualquer sistema de coordenadas.

Solução. Considere o vetor $\mathbf{r}'$ produzido por uma rotação de um ângulo α do vetor $\mathbf{r}$ em torno da direção definida pelo vetor unitário $\hat{\mathbf{n}}$, conforme a Fig. 4.2. Vamos construir uma base ortonormal $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$, indicada na Fig. 4.2, tomando $\hat{\mathbf{e}}_3$ ao longo do eixo de rotação,

$$\hat{\mathbf{e}}_3 = \hat{\mathbf{n}} \ , \tag{4.2.11}$$

e escolhendo $\hat{\mathbf{e}}_1$ de modo que o vetor $\mathbf{r}$ pertença ao plano definido por $\hat{\mathbf{e}}_1$ e $\hat{\mathbf{e}}_3$, sendo $\hat{\mathbf{e}}_2$ perpendicular a este plano (Pearlman 1967). Claramente,

$$\hat{\mathbf{e}}_2 = \frac{\hat{\mathbf{e}}_3 \times \mathbf{r}}{|\hat{\mathbf{e}}_3 \times \mathbf{r}|} = \frac{\hat{\mathbf{n}} \times \mathbf{r}}{r \operatorname{sen} \beta} \ , \tag{4.2.12}$$

onde β é o ângulo entre $\mathbf{r}$ e o eixo de rotação. Além disso,

$$\hat{\mathbf{e}}_1 = \hat{\mathbf{e}}_2 \times \hat{\mathbf{e}}_3 = \frac{\hat{\mathbf{n}} \times (\mathbf{r} \times \hat{\mathbf{n}})}{r \operatorname{sen} \beta} \ . \tag{4.2.13}$$

As componentes x_1', x_2', x_3' de $\mathbf{r}'$ na base $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$ são

$$x_1' = r' \operatorname{sen} \beta \cos \alpha = r \operatorname{sen} \beta \cos \alpha \ , \tag{4.2.14a}$$

$$x_2' = r' \operatorname{sen} \beta \operatorname{sen} \alpha = r \operatorname{sen} \beta \operatorname{sen} \alpha \ , \tag{4.2.14b}$$

$$x_3' = r' \cos \beta = r \cos \beta \ , \tag{4.2.14c}$$

onde usamos $r' = r$. Portanto,

$$\mathbf{r}' = x_1' \hat{\mathbf{e}}_1 + x_2' \hat{\mathbf{e}}_2 + x_3' \hat{\mathbf{e}}_3 = \cos \alpha \, \hat{\mathbf{n}} \times (\mathbf{r} \times \hat{\mathbf{n}}) + \operatorname{sen} \alpha \, \hat{\mathbf{n}} \times \mathbf{r} + (\mathbf{r} \cdot \hat{\mathbf{n}}) \, \hat{\mathbf{n}} \ . \tag{4.2.15}$$

A identidade $\hat{\mathbf{n}} \times (\mathbf{r} \times \hat{\mathbf{n}}) = \mathbf{r} - (\mathbf{r} \cdot \hat{\mathbf{n}}) \, \hat{\mathbf{n}}$ nos permite escrever a *fórmula de rotação*

$$\mathbf{r}' = \cos \alpha \, \mathbf{r} + (1 - \cos \alpha) \, \hat{\mathbf{n}} \, (\hat{\mathbf{n}} \cdot \mathbf{r}) + \operatorname{sen} \alpha \, \hat{\mathbf{n}} \times \mathbf{r} \ . \tag{4.2.16}$$

Introduzindo o operador de rotação $\overleftrightarrow{\mathbf{R}}_{\hat{\mathbf{n}}}(\alpha)$ definido por

$$\overleftrightarrow{\mathbf{R}}_{\hat{\mathbf{n}}}(\alpha) = \cos \alpha \, \overleftrightarrow{\mathbf{1}} + (1 - \cos \alpha) \, \hat{\mathbf{n}} \, \hat{\mathbf{n}} + \operatorname{sen} \alpha \, \hat{\mathbf{n}} \times \overleftrightarrow{\mathbf{1}} \ , \tag{4.2.17}$$

a fórmula de rotação (4.2.16) pode ser sintetizada em

$$\mathbf{r}' = \overleftrightarrow{\mathbf{R}}_{\hat{\mathbf{n}}}(\alpha) \cdot \mathbf{r} \ . \tag{4.2.18}$$

A expressão (4.2.17) é uma representação intrínseca do operador de rotação que só faz referência ao eixo de rotação ($\hat{\mathbf{n}}$) e ao ângulo de rotação (α), sem aludir a nenhum sistema de coordenadas

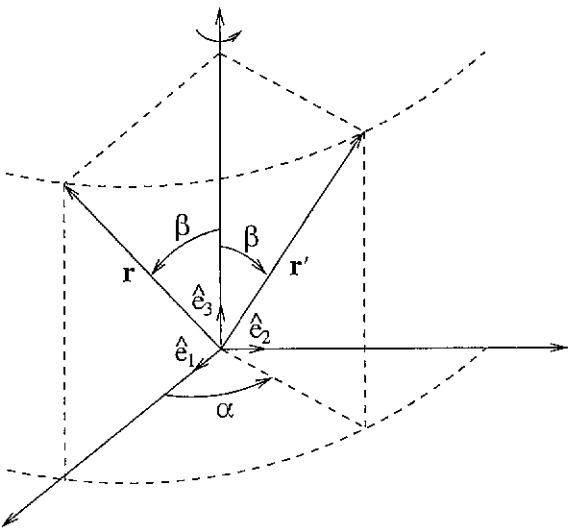


Figura 4.2: Rotação finita de um vetor.

particular. Essa representação é consideravelmente vantajosa em diversas situações de interesse físico (Leubner 1979). ■

■ **Exercício 4.2.1.** Se $\alpha = \delta\theta$ é um ângulo infinitesimal, prove que (4.2.16) reproduz a equação (2.5.10). ■

Lei de Transformação das Componentes de um Tensor

É importante determinar a lei de transformação das componentes de um diádico sob uma rotação dos eixos coordenados. Pela Eq.(3.1.16) temos

$$T'_{rs} = \hat{\mathbf{e}}'_r \cdot \overleftrightarrow{\mathbf{T}} \cdot \hat{\mathbf{e}}'_s = (\sum_k a_{rk}\hat{\mathbf{e}}_k)\cdot \overleftrightarrow{\mathbf{T}} \cdot (\sum_l a_{sl}\hat{\mathbf{e}}_l) = \sum_{k,l} a_{rk}a_{sl}T_{kl} \ . \tag{4.2.19}$$

Um tensor de ordem n ou de n-ésima ordem é definido como um conjunto de 3^n quantidades $T_{k_1\,k_2\,\ldots\,k_n}$ que se transformam sob uma rotação dos eixos coordenados segundo a lei

$$T'_{k_1\,k_2\,\ldots\,k_n} = \sum_{l_1,\,l_2,\ldots,\,l_n} a_{k_1\,l_1}\,a_{k_2\,l_2}\ldots a_{k_n\,l_n}\,T_{l_1\,l_2\,\ldots\,l_n} \ . \tag{4.2.20}$$

Portanto, um diádico é um tensor de segunda ordem. Um vetor é um tensor de primeira ordem e um escalar pode ser considerado um tensor de ordem zero.

Em termos das matrizes $\mathcal{T}' = (T'_{rs})$, $\mathcal{T} = (T_{rs})$ e da matriz de rotação $\mathcal{A}$, a Eq.(4.2.19) é equivalente a

$$T'_{rs} = \sum_k (\mathcal{A})_{rk} \sum_l T_{kl}(\mathcal{A}^T)_{ls} = \sum_k (\mathcal{A})_{rk}(\mathcal{T}\mathcal{A}^T)_{ks} = (\mathcal{A}\mathcal{T}\mathcal{A}^T)_{rs} \ ,$$

isto é,

$$\mathcal{T}' = \mathcal{A}\mathcal{T}\mathcal{A}^{-1} \ . \tag{4.2.21}$$

Portanto, a matriz associada ao diádico $\overleftrightarrow{\mathbf{T}}$ na base $\{\hat{\mathbf{e}}'_i\}_{i=1}^3$ é obtida a partir da matriz associada ao diádico $\overleftrightarrow{\mathbf{T}}$ na base $\{\hat{\mathbf{e}}_i\}_{i=1}^3$ por meio de uma transformação de similaridade executada pela matriz de rotação $\mathcal{A}$ que leva a base $\{\hat{\mathbf{e}}_i\}$ na base $\{\hat{\mathbf{e}}'_i\}$.

Tensores na Física

Tensores de segunda ordem aparecem naturalmente sempre que há uma relação linear entre duas grandezas vetoriais ou quando uma quantidade escalar depende quadraticamente de uma grandeza vetorial. Por exemplo, o tensor de condutividade relaciona a densidade de corrente com o campo elétrico num condutor que obedece à lei de Ohm, e o tensor de permeabilidade magnética estabelece a conexão entre $\mathbf{B}$ e $\mathbf{H}$ num meio magneticamente anisotrópico. Na teoria da elasticidade, a variação do quadrado da distância entre dois pontos infinitesimalmente próximos é quadrática no vetor deslocamento e define o tensor de deformação. Uma grandeza vetorial que dependa bilinearmente de outras duas quantidades vetoriais define um tensor de terceira ordem, e assim por diante. A representação do fluxo de certas grandezas físicas também dá origem a tensores. O fluxo de uma grandeza escalar é descrito por um vetor. Por exemplo, na eletrodinâmica o vetor densidade de corrente $\mathbf{j}$ representa a taxa de transporte de carga elétrica por unidade de área e unidade de tempo. Analogamente, o fluxo de uma grandeza vetorial é expresso por um tensor de segunda ordem. Assim, no eletromagnetismo o tensor de Maxwell T_{ij} descreve o fluxo do momento linear transportado pelo campo eletromagnético. A necessidade de dois índices para caracterizar o fluxo de uma grandeza vetorial é fácil de entender: T_{ij} é o fluxo na j-ésima direção da i-ésima componente do momento linear do campo eletromagnético. A representação do fluxo de uma grandeza descrita por um tensor de segunda ordem requer um tensor de terceira ordem, e assim sucessivamente.

4.3 Momentos e Produtos de Inércia

Em notação diádica, a Eq.(4.1.11) escreve-se

$$\mathbf{L} = \sum_k m_k \left(r_k^2 \overleftrightarrow{\mathbf{1}} - \mathbf{r}_k \mathbf{r}_k\right) \cdot \boldsymbol{\omega} \equiv \overleftrightarrow{\mathbf{I}} \cdot \boldsymbol{\omega} \ , \tag{4.3.1}$$

onde

$$\overleftrightarrow{\mathbf{I}} = \sum_k m_k \left(r_k^2 \overleftrightarrow{\mathbf{1}} - \mathbf{r}_k \mathbf{r}_k\right) \tag{4.3.2}$$

é o *tensor de inércia*, cuja versão contínua é

$$\overleftrightarrow{\mathbf{I}} = \int_V \left(r^2 \overleftrightarrow{\mathbf{1}} - \mathbf{r}\,\mathbf{r}\right) \rho \, dv \ . \tag{4.3.3}$$

As componentes do tensor de inércia na base $\{\hat{\mathbf{e}}_i\}_{i=1}^3$ são dadas pela Eq.(4.1.10), da qual é evidente que o tensor de inércia é simétrico, isto é,

$$I_{ij} = I_{ji} \ , \tag{4.3.4}$$

esta propriedade sendo válida em qualquer base. A matriz cujos elementos são as componentes do tensor de inércia relativamente a um dado sistema de coordenadas cartesianas é conhecida como *matriz de inércia*, denotada por I. Os elementos da diagonal da *matriz de inércia* $\mathrm{I} = (I_{ij})$ são quantidades necessariamente positivas chamadas de *momentos de inércia*:

$$I_{11} \equiv I_{xx} = \int_V (r^2 - x^2)\,\rho\,dv = \int_V (y^2 + z^2)\,\rho\,dv \ ; \tag{4.3.5a}$$

$$I_{22} \equiv I_{yy} = \int_V (r^2 - y^2)\,\rho\,dv = \int_V (x^2 + z^2)\,\rho\,dv \ ; \tag{4.3.5b}$$

$$I_{33} \equiv I_{zz} = \int_V (r^2 - z^2)\,\rho\,dv = \int_V (x^2 + y^2)\,\rho\,dv \ . \tag{4.3.5c}$$

Os elementos de fora da diagonal da matriz de inércia são chamados de *produtos de inércia* e são dados por:

$$I_{12} = I_{21} \equiv I_{yx} = I_{xy} = -\int_V xy\,\rho\,dv \ ; \tag{4.3.6a}$$

$$I_{13} = I_{31} \equiv I_{zx} = I_{xz} = -\int_V xz\,\rho\,dv \ ; \tag{4.3.6b}$$

$$I_{23} = I_{32} \equiv I_{zy} = I_{yz} = -\int_V yz\,\rho\,dv \ . \tag{4.3.6c}$$

É importante sublinhar que os momentos e produtos de inércia dependem não apenas das direções dos eixos coordenados escolhidos, mas também de sua origem.

4.4 Energia Cinética e Teorema dos Eixos Paralelos

A energia cinética de rotação em torno do ponto O (Fig. 4.1) é dada por

$$T = \sum_k \frac{m_k}{2} v_k^2 = \sum_k \frac{m_k}{2} \mathbf{v}_k \cdot (\boldsymbol{\omega} \times \mathbf{r}_k) = \sum_k \frac{m_k}{2} \boldsymbol{\omega} \cdot (\mathbf{r}_k \times \mathbf{v}_k) = \frac{1}{2}\boldsymbol{\omega} \cdot \sum_k \mathbf{r}_k \times \mathbf{p}_k = \frac{1}{2}\boldsymbol{\omega} \cdot \mathbf{L} \ . \tag{4.4.1}$$

Com o emprego da Eq.(4.3.1) podemos escrever

$$T = \frac{1}{2}\boldsymbol{\omega} \cdot \overleftrightarrow{\mathbf{I}} \cdot \boldsymbol{\omega} \ . \tag{4.4.2}$$

Seja $\hat{\mathbf{n}}$ o unitário ao longo do eixo instantâneo de rotação, de modo que $\boldsymbol{\omega} = \omega\hat{\mathbf{n}}$. Então,

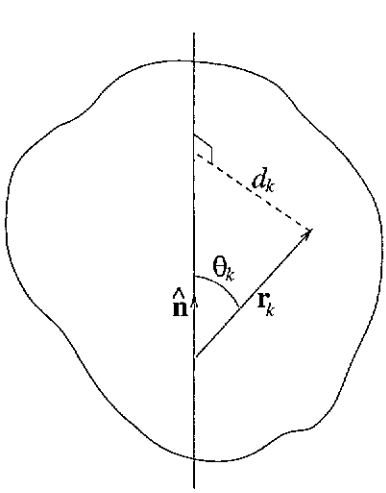


Figura 4.3: Momento de inércia: d_k é a distância da k-ésima partícula ao eixo de rotação.

$$T = \frac{1}{2}\hat{\mathbf{n}}\cdot\overleftrightarrow{\mathbf{I}}\cdot\hat{\mathbf{n}}\,\omega^2 \equiv \frac{1}{2}I\omega^2 \ , \tag{4.4.3}$$

onde

$$I = \hat{\mathbf{n}}\cdot\overleftrightarrow{\mathbf{I}}\cdot\hat{\mathbf{n}} \tag{4.4.4}$$

é chamado de momento de inércia em relação ao eixo de rotação. Mais explicitamente, com o uso de (4.3.2),

$$I = \hat{\mathbf{n}}\cdot\overleftrightarrow{\mathbf{I}}\cdot\hat{\mathbf{n}} = \sum_k m_k[r_k^2 - (\hat{\mathbf{n}}\cdot\mathbf{r}_k)^2] = \sum_k m_k d_k^2 \tag{4.4.5}$$

onde

$$d_k^2 = r_k^2 - (\hat{\mathbf{n}}\cdot\mathbf{r}_k)^2 = r_k^2(1-\cos^2\theta_k) = r_k^2\,\mathrm{sen}^2\,\theta_k = (\hat{\mathbf{n}}\times\mathbf{r}_k)^2 \tag{4.4.6}$$

é o quadrado da distância da k-ésima partícula ao eixo de rotação (Fig. 4.3). A expressão (4.4.5) coincide com a definição de momento de inércia encontrada nos textos de física elementar. Da mesma forma, (4.4.3) é a fórmula usual para a energia cinética de rotação de um corpo rígido. Em geral, no entanto, a velocidade angular muda de direção com o passar do tempo, de modo que I varia com o tempo. O momento de inércia I é constante se o corpo estiver limitado a girar em torno de um eixo fixo, caso normalmente tratado na física elementar.

O momento de inércia depende da escolha da origem do sistema de coordenadas adotado para calculá-lo. Há, todavia, uma relação simples entre momentos de inércia relativos a eixos paralelos quando um deles passa pelo centro de massa.

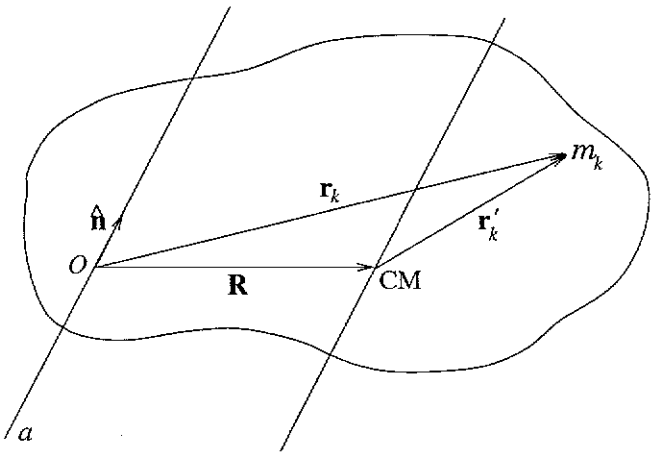


Figura 4.4: Teorema dos eixos paralelos.

Teorema dos Eixos Paralelos.[3] O momento de inércia em relação a um dado eixo é igual ao momento de inércia relativo a um eixo paralelo passando pelo centro de massa acrescido do momento de inércia do corpo em relação ao eixo original, calculado como se o corpo estivesse inteiramente concentrado no centro de massa.

Demonstração. Da Fig. 4.4 infere-se $\mathbf{r}_k = \mathbf{R} + \mathbf{r}'_k$ onde $\mathbf{R}$ é o vetor posição do centro de massa a contar da origem O. O momento de inércia em relação ao eixo a é, por (4.4.5) e (4.4.6),

$$I_a = \sum_k m_k(\hat{\mathbf{n}} \times \mathbf{r}_k)^2 = \sum_k m_k(\hat{\mathbf{n}} \times \mathbf{R} + \hat{\mathbf{n}} \times \mathbf{r}'_k)^2 =$$

$$= (\sum_k m_k)(\hat{\mathbf{n}} \times \mathbf{R})^2 + \sum_k m_k(\hat{\mathbf{n}} \times \mathbf{r}'_k)^2 + 2(\hat{\mathbf{n}} \times \mathbf{R})\cdot(\hat{\mathbf{n}} \times \sum_k m_k \mathbf{r}'_k) \ . \qquad (4.4.7)$$

Com o emprego da Eq.(1.1.18) esta última equação reduz-se a

$$I_a = I_{CM} + M(\hat{\mathbf{n}} \times \mathbf{R})^2 \ , \qquad (4.4.8)$$

completando a demonstração do teorema, pois, conforme a Eq.(4.4.6), $(\hat{\mathbf{n}} \times \mathbf{R})^2$ vem a ser o quadrado da distância do centro de massa ao eixo original a. □

O próprio tensor de inércia admite uma decomposição análoga à (4.4.8). De fato, substituindo $\mathbf{r}_k = \mathbf{R} + \mathbf{r}'_k$ em (4.3.2) e utilizando as propriedades do produto tensorial, resulta

$$\overleftrightarrow{\mathbf{I}}_O = \sum_k m_k[(r'^2_k + 2\mathbf{R}\cdot\mathbf{r}'_k + R^2)\overleftrightarrow{\mathbf{1}} - (\mathbf{R} + \mathbf{r}'_k)(\mathbf{R} + \mathbf{r}'_k)]$$

$$= \sum_k m_k(r'^2_k \overleftrightarrow{\mathbf{1}} - \mathbf{r}'_k\mathbf{r}'_k) + (R^2\overleftrightarrow{\mathbf{1}} - \mathbf{RR})\sum_k m_k$$

[3] Também conhecido como *teorema de Steiner*.

$$+2\mathbf{R}\cdot(\sum_k m_k\mathbf{r}'_k)\overleftrightarrow{1} - \mathbf{R}(\sum_k m_k\mathbf{r}'_k) - (\sum_k m_k\mathbf{r}'_k)\mathbf{R}$$

$$= \sum_k m_k({r'_k}^2\overleftrightarrow{1} - \mathbf{r}'_k\mathbf{r}'_k) + M(R^2\overleftrightarrow{1} - \mathbf{RR}) \ ,$$

isto é,

$$\overleftrightarrow{\mathbf{I}}_O = \overleftrightarrow{\mathbf{I}}_{CM} + M(R^2\overleftrightarrow{1} - \mathbf{RR}) \ . \tag{4.4.9}$$

Note que o segundo termo é o tensor de inércia, relativo à origem O, do corpo rígido como se toda a sua massa estivesse concentrada no centro de massa.

■ **Exercício 4.4.1.** Deduza (4.4.8) tomando (4.4.9) como ponto de partida. ■

Teorema do Eixo Perpendicular. Dada uma placa plana de formato e distribuição de massa arbitrários, a soma de seus momentos de inércia em relação a quaisquer dois eixos perpendiculares contidos no plano da placa é igual ao momento de inércia relativamente a um eixo que passa por seu ponto de interseção e é perpendicular à placa.

■ **Exercício 4.4.2.** Demonstre o teorema do eixo perpendicular . ■

4.5 Diagonalização do Tensor de Inércia

As componentes do tensor de inércia dependem da origem e da orientação dos eixos coordenados relativamente ao corpo rígido. Uma vez fixada uma origem, seria extremamente conveniente se fosse possível encontrar eixos coordenados ξ_1, ξ_2, ξ_3 com vetores unitários $\hat{\boldsymbol{\xi}}_1, \hat{\boldsymbol{\xi}}_2, \hat{\boldsymbol{\xi}}_3$ relativamente aos quais o tensor de inércia assumisse a forma diagonal. Nesse caso teríamos

$$\overleftrightarrow{\mathbf{I}} = I_1\,\hat{\boldsymbol{\xi}}_1\hat{\boldsymbol{\xi}}_1 + I_2\,\hat{\boldsymbol{\xi}}_2\hat{\boldsymbol{\xi}}_2 + I_3\,\hat{\boldsymbol{\xi}}_3\hat{\boldsymbol{\xi}}_3 \tag{4.5.1}$$

donde

$$\mathbf{L} = \overleftrightarrow{\mathbf{I}}\cdot\boldsymbol{\omega} = I_1\omega_1\,\hat{\boldsymbol{\xi}}_1 + I_2\omega_2\,\hat{\boldsymbol{\xi}}_2 + I_3\omega_3\,\hat{\boldsymbol{\xi}}_3 \ , \tag{4.5.2}$$

isto é,

$$L_1 = I_1\omega_1 \quad , \quad L_2 = I_2\omega_2 \quad , \quad L_3 = I_3\omega_3 \ . \tag{4.5.3}$$

Analogamente, a energia cinética tomaria a forma particularmente simples

$$T = \frac{1}{2}\omega\cdot\overleftrightarrow{\mathbf{I}}\cdot\omega = \frac{1}{2}\omega\cdot\mathbf{L} = \frac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2) \ . \tag{4.5.4}$$

O tensor de inércia é um operador linear real e simétrico, logo auto-adjunto, e o Teorema C.13 do Apêndice C assegura a existência de uma base ortonormal constituída por autovetores de $\overleftrightarrow{\mathbf{I}}$. Em outras palavras, existem três números reais positivos I_1, I_2, I_3 e três vetores unitários e mutuamente ortogonais $\hat{\boldsymbol{\xi}}_1$, $\hat{\boldsymbol{\xi}}_2$, $\hat{\boldsymbol{\xi}}_3$ tais que

$$\overleftrightarrow{\mathbf{I}}\cdot\hat{\boldsymbol{\xi}}_1 = I_1\hat{\boldsymbol{\xi}}_1 \quad , \quad \overleftrightarrow{\mathbf{I}}\cdot\hat{\boldsymbol{\xi}}_2 = I_2\hat{\boldsymbol{\xi}}_2 \quad , \quad \overleftrightarrow{\mathbf{I}}\cdot\hat{\boldsymbol{\xi}}_3 = I_3\hat{\boldsymbol{\xi}}_3 \tag{4.5.5}$$

ou, em forma abreviada,

$$\overleftrightarrow{\mathbf{I}}\cdot\hat{\boldsymbol{\xi}}_j = I_j\hat{\boldsymbol{\xi}}_j \quad , \quad j = 1,2,3 \ . \tag{4.5.6}$$

Escolhendo os eixos cartesianos ao longo dos vetores $\hat{\boldsymbol{\xi}}_1$, $\hat{\boldsymbol{\xi}}_2$, $\hat{\boldsymbol{\xi}}_3$, as componentes de $\overleftrightarrow{\mathbf{I}}$ nesta base são dadas por

$$I_{ij} = \hat{\boldsymbol{\xi}}_i\cdot\overleftrightarrow{\mathbf{I}}\cdot\hat{\boldsymbol{\xi}}_j = I_j\,\hat{\boldsymbol{\xi}}_i\cdot\hat{\boldsymbol{\xi}}_j = I_j\,\delta_{ij} \ , \tag{4.5.7}$$

de modo que a matriz de inércia nesta base é diagonal:

$$\mathbf{I} = (I_{ij}) = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \ . \tag{4.5.8}$$

Os elementos da diagonal, I_1 , I_2 e I_3 , são chamados de *momentos principais de inércia*, e os eixos cartesianos correspondentes são chamados de *eixos principais de inércia*.

A determinação dos eixos principais de inércia equivale à resolução da equação para $\boldsymbol{\xi}$

$$\overleftrightarrow{\mathbf{I}}\cdot\boldsymbol{\xi} = I\,\boldsymbol{\xi} \ , \tag{4.5.9}$$

com $\boldsymbol{\xi} \neq 0$ e $I \in \mathbb{R}$. Encontradas três soluções mutuamente ortogonais ficam definidos os eixos principais de inércia, os valores correspondentes de I sendo os momentos principais de inércia. Com uma escolha inicial arbitrária dos eixos coordenados, a Eq.(4.5.9) escreve-se, em componentes,

$$\sum_{j=1}^{3} I_{ij}\,\xi_j = I\,\xi_i \quad , \quad i = 1,2,3 \ , \tag{4.5.10}$$

ou, matricialmente,

$$\begin{pmatrix} I_{11} - I & I_{12} & I_{13} \\ I_{21} & I_{22} - I & I_{23} \\ I_{31} & I_{32} & I_{33} - I \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = 0 \ . \tag{4.5.11}$$

Este sistema de equações lineares para as componentes de $\boldsymbol{\xi}$ só tem solução não trivial se

$$\det \begin{pmatrix} I_{11}-I & I_{12} & I_{13} \\ I_{21} & I_{22}-I & I_{23} \\ I_{31} & I_{32} & I_{33}-I \end{pmatrix} = 0 \ . \tag{4.5.12}$$

Esta última equação é cúbica em I, e suas três raízes são os momentos principais de inércia, isto é, são os autovalores do tensor de inércia. Para cada autovalor encontrado, a Eq.(4.5.11) pode ser resolvida para as componentes do autovetor associado $\boldsymbol{\xi}$, fornecendo, assim, a direção do eixo principal de inércia correspondente.

■ **Exemplo 4.5.1.** Dada uma placa homogênea na forma de um triângulo retângulo isósceles, determinar os eixos principais de inércia em relação ao vértice oposto à hipotenusa.

Solução. Adotamos inicialmente o sistema de eixos indicado na Fig. 4.5(a). Se M é a massa da placa, sua densidade superficial de massa é

$$\sigma = \frac{M}{a^2/2} = \frac{2M}{a^2} \ .$$

O tensor de inércia tem as seguintes componentes:

$$I_{11} \equiv I_{xx} = \sigma \int (y^2+z^2)da = \sigma \int_0^a \int_0^{a-y} y^2 dxdy = \frac{Ma^2}{6} \ ;$$

$$I_{22} \equiv I_{yy} = \sigma \int (x^2+z^2)da = I_{11} = \frac{Ma^2}{6} \quad \text{(por simetria)} \ ;$$

$$I_{33} \equiv I_{zz} = I_{xx} + I_{yy} = \frac{Ma^2}{3} \quad \text{(pelo teorema do eixo perpendicular)} \ ;$$

$$I_{12} \equiv I_{xy} = -\sigma \int xyda = -\sigma \int_0^a \int_0^{a-y} xydxdy = -\frac{Ma^2}{12} \ ;$$

$$I_{13} \equiv I_{xz} = -\sigma \int xzda = 0 \ ; \ I_{23} \equiv I_{yz} = -\sigma \int yzda = 0 \ .$$

Portanto,

$$\mathrm{I} = (I_{ij}) = \frac{Ma^2}{12} \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \ .$$

Pondo $I = Ma^2\lambda/12$, precisamos encontrar as raízes da equação

$$\det \begin{pmatrix} 2-\lambda & -1 & 0 \\ -1 & 2-\lambda & 0 \\ 0 & 0 & 4-\lambda \end{pmatrix} = (4-\lambda)[(2-\lambda)^2-1] = 0 \ .$$

As raízes são $\lambda_1 = 1, \lambda_2 = 3, \lambda_3 = 4$, de modo que os momentos principais de inércia são

$$I_1 = \frac{Ma^2}{12} \ , \ I_2 = \frac{Ma^2}{4} \ , \ I_3 = \frac{Ma^2}{3} \ .$$

O autovetor com autovalor I_1 satisfaz

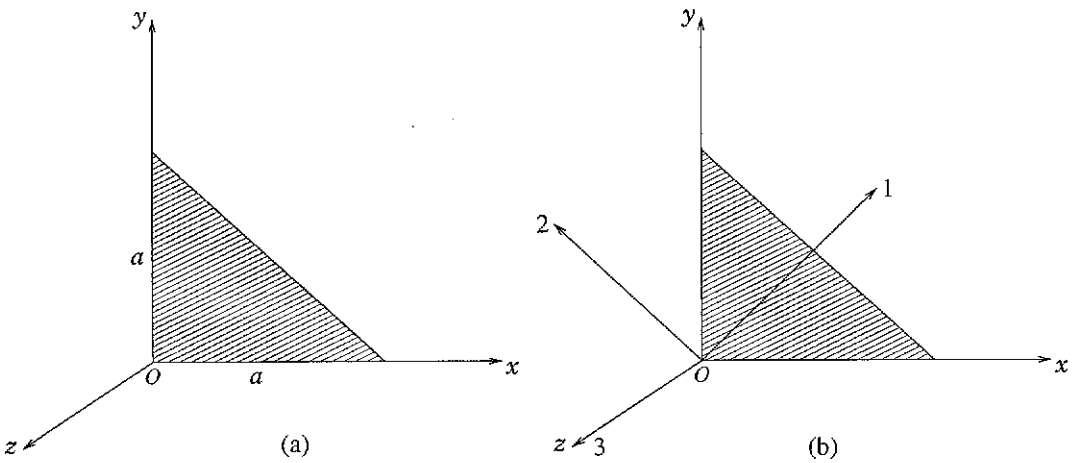


Figura 4.5: Eixos originalmente escolhidos e eixos principais de inércia de uma placa plana triangular em relação ao vértice oposto à hipotenusa.

$$\begin{pmatrix} 2-\lambda_1 & -1 & 0 \\ -1 & 2-\lambda_1 & 0 \\ 0 & 0 & 4-\lambda_1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = 0 \implies \begin{cases} \xi_1 - \xi_2 = 0 \\ -\xi_1 + \xi_2 = 0 \\ 3\xi_3 = 0 \end{cases} .$$

Conseqüentemente,

$$\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{ou} \quad \boldsymbol{\xi} = \alpha(\hat{\mathbf{i}} + \hat{\mathbf{j}}) \ ,$$

onde α é um número real arbitrário. Escolhendo-se $\alpha = 1/\sqrt{2}$ resulta um vetor normalizado (unitário) $\hat{\boldsymbol{\xi}}_1$. Este vetor, que pertence ao plano xy e faz um ângulo de 45^o com o eixo x, fornece a direção do primeiro eixo principal de inércia. Repetindo o procedimento, obtém-se o segundo eixo principal de inércia ao longo do vetor $\hat{\boldsymbol{\xi}}_2 = 2^{-1/2}(-\hat{\mathbf{i}} + \hat{\mathbf{j}})$, e o terceiro ao longo do vetor $\boldsymbol{\xi}_3 = \hat{\mathbf{k}}$, isto é, coincidente com o eixo z [Fig. 4.5(b)]. ■

■ **Exemplo 4.5.2.** Usando a equação (4.4.9), mostrar que os eixos ξ_1, ξ_2, ξ_3 da Fig. 4.5(b) também são eixos principais de inércia em relação ao centro de massa da placa e determine os momentos principais de inércia correspondentes.

Solução. O vetor posição do centro de massa a contar da origem O é $\mathbf{R} = (a\sqrt{2}/3)\boldsymbol{\xi}_1$, donde, por (4.4.9),

$$\overleftrightarrow{\mathbf{I}}_{CM} = \overleftrightarrow{\mathbf{I}}_O - M(R^2\,\overleftrightarrow{\mathbf{1}} - \mathbf{RR}) = I_1\hat{\boldsymbol{\xi}}_1\hat{\boldsymbol{\xi}}_1 + I_2\hat{\boldsymbol{\xi}}_2\hat{\boldsymbol{\xi}}_2 + I_3\hat{\boldsymbol{\xi}}_3\hat{\boldsymbol{\xi}}_3 - \frac{2Ma^2}{9}(\hat{\boldsymbol{\xi}}_1\hat{\boldsymbol{\xi}}_1 + \hat{\boldsymbol{\xi}}_2\hat{\boldsymbol{\xi}}_2 + \hat{\boldsymbol{\xi}}_3\hat{\boldsymbol{\xi}}_3 - \hat{\boldsymbol{\xi}}_1\hat{\boldsymbol{\xi}}_1)$$

$$= I_1\hat{\boldsymbol{\xi}}_1\hat{\boldsymbol{\xi}}_1 + (I_2 - \frac{2Ma^2}{9})\hat{\boldsymbol{\xi}}_2\hat{\boldsymbol{\xi}}_2 + (I_3 - \frac{2Ma^2}{9})\hat{\boldsymbol{\xi}}_3\hat{\boldsymbol{\xi}}_3 \ .$$

Logo, $\overleftrightarrow{\mathbf{I}}_{CM}$ também é diagonal relativamente aos eixos ξ_1, ξ_2, ξ_3 com origem no centro de massa e

$$I_1^{\rm CM} = \frac{Ma^2}{12} \quad , \quad I_2^{\rm CM} = \frac{Ma^2}{36} \quad , \quad I_3^{\rm CM} = \frac{Ma^2}{9}$$

são os momentos principais de inércia correspondentes. ■

4.6 Simetrias e Eixos Principais de Inércia

A identificação dos eixos principais de inércia é grandemente facilitada se o corpo é simétrico. Dizemos que um corpo rígido possui um plano de simetria se existe um plano que divide o corpo em duas porções de tal modo que uma é a imagem especular da outra relativamente ao referido plano. Em particular, a densidade de massa do corpo deve ser a mesma em pontos simetricamente situados em relação ao plano de simetria.

Lema 4.6.1. Se um corpo tem um plano de simetria passando pelo ponto em relação ao qual o tensor de inércia está referido, um eixo principal de inércia é perpendicular a esse plano.

Demonstração. Seja xy o plano de simetria, de modo que $\rho(x,y,-z) = \rho(x,y,z)$. Conseqüentemente,

$$I_{xz} = -\int\int\int xz\rho(x,y,z)dxdydz = 0 \tag{4.6.1}$$

porque o integrando é uma função ímpar de z e o intervalo de integração é simétrico em relação a $z = 0$. Argumento análogo aplica-se a I_{yz}, de modo que

$$\mathbf{I} = \begin{pmatrix} I_{xx} & I_{xy} & 0 \\ I_{yx} & I_{yy} & 0 \\ 0 & 0 & I_{zz} \end{pmatrix} , \tag{4.6.2}$$

que claramente possui o autovetor

$$\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{4.6.3}$$

ou $\boldsymbol{\xi} = \hat{\mathbf{k}}$ com autovalor I_{zz}. □

Diz-se que um corpo possui um eixo de simetria se ele é uma figura de revolução em torno do referido eixo e a distribuição de massa não depende do ângulo de rotação em torno do eixo.

Lema 4.6.2. Um eixo de simetria de um corpo rígido passando pelo ponto em relação ao qual o tensor de inércia está referido é um eixo principal de inércia. Quaisquer dois eixos mutuamente ortogonais contidos num plano perpendicular ao eixo de simetria são eixos principais de inércia, e os momentos principais de inércia correspondentes são iguais entre si.

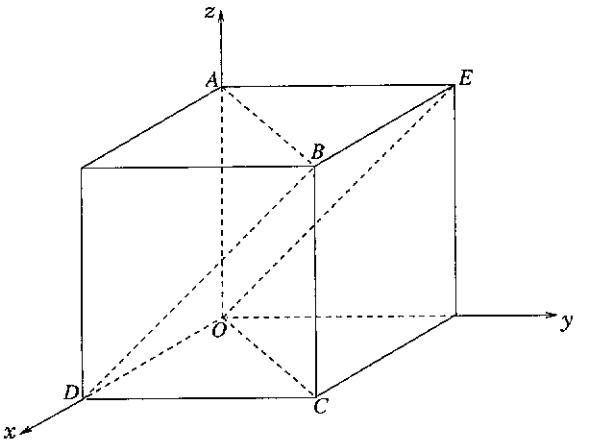


Figura 4.6: Planos de simetria de um cubo homogêneo em relação a um vértice.

■ **Exercício 4.6.1.** Demonstre o Lema 4.6.2. Sugestão: considere dois planos perpendiculares cuja interseção é o eixo de simetria e aplique o Lema 4.6.1 a cada um deles. ■

Um reexame do Exemplo 4.5.1 à luz dos lemas anteriores revela que os eixos principais de inércia da placa triangular poderiam ter sido identificados sem cálculos. O eixo z é perpendicular ao plano da placa (Fig. 4.5), que obviamente é um plano de simetria. Portanto, o Lema 4.6.1 garante que o eixo z é um eixo principal de inércia. Claramente, um plano perpendicular à placa que passa pela origem O e é ortogonal à hipotenusa é também um plano de simetria, de modo que, novamente pelo Lema 4.6.1, o eixo 2 é um eixo principal de inércia. Finalmente, o terceiro eixo é simplesmente perpendicular aos outros dois, e identificamos o eixo 1 como o último eixo principal de inércia da placa.

■ **Exemplo 4.6.1.** Determinar os eixos principais de inércia de um cubo homeogêneo, de massa M e aresta a, em relação a um dos vértices.

Solução. Escolhamos os eixos iniciais conforme indicado na Fig. 4.6. Por simetria, os momentos de inêrcia são iguais entre si, o mesmo acontecendo com os produtos de inércia:

$$I_{zz} = I_{yy} = I_{xx} = \int_0^a \int_0^a \int_0^a (y^2 + z^2)\rho dx dy dz = \frac{2}{3}\rho a^5 = \frac{2}{3}Ma^2 \ ;$$

$$I_{xz} = I_{yz} = I_{xy} = -\int_0^a \int_0^a \int_0^a xy\rho dx dy dz = -\frac{1}{4}\rho a^5 = -\frac{Ma^2}{4} \ .$$

Assim, a matriz de inércia é dada por

$$\mathrm{I} = \frac{2Ma^2}{3}\begin{pmatrix} 1 & -3/8 & -3/8 \\ -3/8 & 1 & -3/8 \\ -3/8 & -3/8 & 1 \end{pmatrix} \ . \tag{4.6.4}$$

Por inspeção da Fig. 4.6 identificamos os planos de simetria passando pelo vértice $OABCO$ e $ODBEO$, cujos vetores unitários normais são, respectivamente,

$$\hat{\mathbf{n}}_1 = \frac{\overrightarrow{OC}}{\sqrt{2}a} \times \frac{\overrightarrow{OA}}{a} = \frac{1}{\sqrt{2}}(\hat{\mathbf{i}}+\hat{\mathbf{j}})\times\hat{\mathbf{k}} = \frac{1}{\sqrt{2}}(\hat{\mathbf{i}}-\hat{\mathbf{j}}) \ , \quad \hat{\mathbf{n}}_2 = \frac{\overrightarrow{OD}}{a} \times \frac{\overrightarrow{OE}}{\sqrt{2}a} = \hat{\mathbf{i}} \times \frac{(\hat{\mathbf{j}}+\hat{\mathbf{k}})}{\sqrt{2}} = \frac{1}{\sqrt{2}}(\hat{\mathbf{k}}-\hat{\mathbf{j}}) \ .$$

Os dois planos de simetria considerados não são perpendiculares, de modo que os vetores $\hat{\mathbf{n}}_1$ e $\hat{\mathbf{n}}_2$ são linearmente independentes mas não são mutuamente ortogonais. Isto é possível porque a matriz de inércia tem autovalores degenerados (vide Apêndice C), isto é, dois dos momentos principais de inércia são iguais. De fato, de

$$\begin{pmatrix} 1 & \alpha & \alpha \\ \alpha & 1 & \alpha \\ \alpha & \alpha & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = (1-\alpha) \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} , \quad \begin{pmatrix} 1 & \alpha & \alpha \\ \alpha & 1 & \alpha \\ \alpha & \alpha & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = (1-\alpha) \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

com $\alpha = -3/8$ deduz-se que $I_1 = I_2 = 11Ma^2/12$. Autovetores mutuamente ortogonais e normalizados podem ser obtidos fazendo, por exemplo, $\hat{\boldsymbol{\xi}}_1 = \hat{\mathbf{n}}_1$ e $\hat{\boldsymbol{\xi}}_2 = \gamma(\,\hat{\mathbf{n}}_1 + \beta\hat{\mathbf{n}}_2\,)$, com os números reais β e γ escolhidos de tal modo que $\boldsymbol{\xi}_1 \cdot \boldsymbol{\xi}_2 = 0$ e $|\hat{\boldsymbol{\xi}}_2| = 1$. Um cálculo elementar fornece $\hat{\boldsymbol{\xi}}_2 = (\hat{\mathbf{i}}+\hat{\mathbf{j}}-2\hat{\mathbf{k}})/\sqrt{6}$. O vetor ao longo do terceiro eixo principal de inércia é simplesmente

$$\hat{\boldsymbol{\xi}}_3 = \hat{\boldsymbol{\xi}}_1 \times \hat{\boldsymbol{\xi}}_2 = \frac{(\hat{\mathbf{i}}-\hat{\mathbf{j}})}{\sqrt{2}} \times \frac{(\hat{\mathbf{i}}+\hat{\mathbf{j}}-2\hat{\mathbf{k}})}{\sqrt{6}} = \frac{1}{\sqrt{3}}(\,\hat{\mathbf{i}}+\hat{\mathbf{j}}+\hat{\mathbf{k}}\,) \ ,$$

cuja direção coincide com a da diagonal $\overline{OB}$ do cubo. Finalmente, de

$$\begin{pmatrix} 1 & -3/8 & -3/8 \\ -3/8 & 1 & -3/8 \\ -3/8 & -3/8 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} ,$$

conclui-se que o terceiro momento principal de inércia é $I_3 = Ma^2/6$. ■

4.7 Moeda Rolante

Mesmo os problemas tridimensionais mais simples da dinâmica dos corpos rígidos já são consideravelmente complicados. No outro extremo situam-se os problemas elementares de rotação em torno de um eixo fixo. Há problemas interessantes com apenas dois graus de liberdade rotacionais, os quais caracterizam-se por um grau de dificuldade intermediário e fornecem ilustrações bastante ricas e instrutivas dos conceitos e técnicas empregados na investigação do movimento dos corpos rígidos.

Consideremos, por exemplo, uma moeda de massa m e raio R que cai ao longo de uma mesa inclinada rolando sem deslizar e sem tombar – o seu plano permanece perpendicular à mesa. Os eixos principais de inércia fixos na moeda e passando pelo centro de massa são quaisquer dois eixos mutuamente ortogonais contidos no plano da moeda, com o terceiro eixo perpendicular ao plano da moeda. Um cálculo elementar fornece $I_3 = mR^2/2$ e, por simetria e pelo teorema do eixo perpendicular, $I_1 = I_2 = I_3/2 = mR^2/4$. A energia cinética da moeda é a energia de translação do centro de massa acrescida da energia de rotação em torno do centro de massa, de modo que a lagrangiana do sistema escreve-se

$$L = T - V = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2) + mgy \operatorname{sen}\alpha \ , \qquad (4.7.1a)$$

onde α é a inclinação da mesa e x, y são coordenadas cartesianas do centro da moeda. Sejam ϕ e θ ângulos que descrevem rotações da moeda em torno do seu eixo de simetria

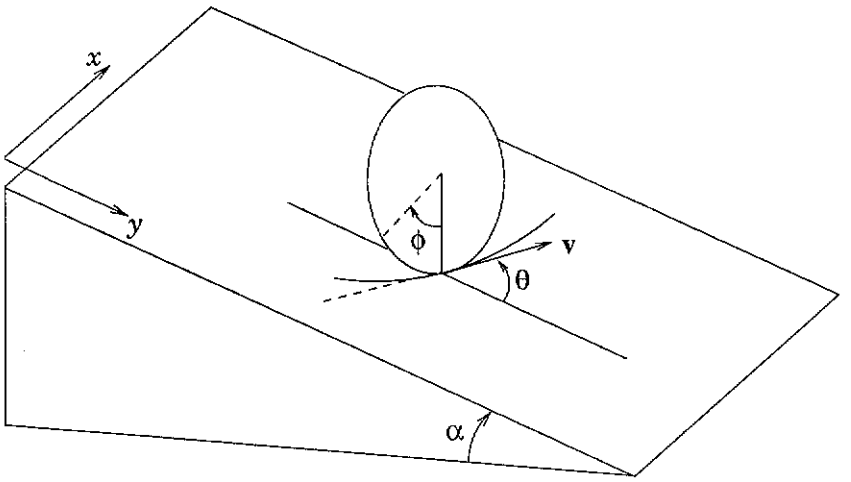


Figura 4.7: Moeda rolando sem deslizar ao longo de um plano inclinado fixo.

e de um eixo perpendicular à mesa, respectivamente, como indicado na Fig. 4.7. Conseqüentemente, $\omega_3 = \dot{\phi}$ e, levando em conta que $\dot{\theta}$ representa a magnitude da projeção do vetor velocidade angular sobre o plano da moeda, $\omega_1^2 + \omega_2^2 = \dot{\theta}^2$. Assim, a lagrangiana do sistema toma a forma

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) + \frac{mR^2}{4}\dot{\phi}^2 + \frac{mR^2}{8}\dot{\theta}^2 + mgy \operatorname{sen}\alpha \ . \qquad (4.7.1b)$$

Os vínculos de rolamento são

$$\dot{x} - R\dot{\phi}\operatorname{sen}\theta = 0 \ , \qquad (4.7.2a)$$

$$\dot{y} - R\dot{\phi}\cos\theta = 0 \ . \qquad (4.7.2b)$$

Lançando mão do método dos multiplicadores de Lagrange, encontramos as equações de movimento

$$m\ddot{x} = \lambda_1 \ , \qquad (4.7.3a)$$

$$m\ddot{y} - mg\operatorname{sen}\alpha = \lambda_2 \ , \qquad (4.7.3b)$$

$$\frac{mR^2}{4}\ddot{\theta} = 0 \ , \qquad (4.7.3c)$$

$$\frac{mR^2}{2}\ddot{\phi} = -\lambda_1 R\operatorname{sen}\theta - \lambda_2 R\cos\theta \ , \qquad (4.7.3d)$$

onde λ_1, λ_2 são os multiplicadores de Lagrange associados aos dois vínculos não-holônomos (4.7.2).

■ **Exercício 4.7.1.** Deduza as equações (4.7.3). ■

De (4.7.3c) resulta imediatamente

$$\theta = \theta_0 + \Omega t \ , \tag{4.7.4}$$

onde θ_0 e Ω são constantes arbitrárias. Combinando as Eqs.(4.7.2a) e (4.7.3a) obtém-se

$$\lambda_1 = mR\ddot{\phi}\,\mathrm{sen}\,\theta + mR\Omega\dot{\phi}\cos\theta \ , \tag{4.7.5}$$

e, procedendo analogamente,

$$\lambda_2 = mR\ddot{\phi}\cos\theta - mR\Omega\dot{\phi}\,\mathrm{sen}\,\theta - mg\,\mathrm{sen}\,\alpha \ . \tag{4.7.6}$$

A substituição dessas expressões para λ_1 e λ_2 em (4.7.3d) conduz à seguinte equação diferencial para ϕ:

$$\ddot{\phi} = \frac{2g\,\mathrm{sen}\,\alpha}{3R}\cos(\theta_0 + \Omega t) \ . \tag{4.7.7}$$

A solução geral desta equação é

$$\phi = \phi_0 + \omega t - \frac{2g\,\mathrm{sen}\,\alpha}{3\Omega^2 R}\cos(\theta_0 + \Omega t) \ . \tag{4.7.8}$$

Finalmente, levando as Eqs.(4.7.4) e (4.7.8) nas equações de vínculo (4.7.2) e integrando resultam

$$x = x_0 + \frac{g\,\mathrm{sen}\,\alpha}{3\Omega}t - \left[\frac{\omega R}{\Omega} + \frac{g\,\mathrm{sen}\,\alpha}{3\Omega^2}\,\mathrm{sen}\,(\theta_0 + \Omega t)\right]\cos(\theta_0 + \Omega t) \ , \tag{4.7.9}$$

$$y = y_0 + \left[\frac{\omega R}{\Omega} + \frac{g\,\mathrm{sen}\,\alpha}{3\Omega^2}\,\mathrm{sen}\,(\theta_0 + \Omega t)\right]\mathrm{sen}\,(\theta_0 + \Omega t) \ . \tag{4.7.10}$$

Isto completa a integração das equações de movimento da moeda rolante em termos das seis constantes arbitrárias $x_0, y_0, \theta_0, \phi_0, \omega, \Omega$. Surpreendentemente, o acoplamento entre a translação e a rotação faz com que o movimento do centro de massa da moeda ao longo do eixo y seja limitado e oscilante: se o plano inclinado for suficientemente extenso, a moeda sempre permanecerá sobre o plano.

■ **Exercício 4.7.2.** A solução que acabamos de encontrar para o movimento da moeda rolante não vale se $\Omega = 0$. Obtenha a solução neste caso. ■

4.8 Equações de Euler e Rotação Livre

Seja $\mathbf{N}$ o torque em relação a um ponto O, que é um ponto fixo do corpo rígido ou é o seu centro de massa. Do ponto de vista de uma sistema inercial temos

$$\left(\frac{d\mathbf{L}}{dt}\right)_{inercial} = \mathbf{N} \ , \tag{4.8.1}$$

donde

$$\left(\frac{d\mathbf{L}}{dt}\right)_{corpo} + \boldsymbol{\omega} \times \mathbf{L} = \mathbf{N} . \tag{4.8.2}$$

O tensor de inércia varia com o tempo se estiver referido a eixos fixos no espaço. Relativamente a eixos atados ao corpo, no entanto, o tensor de inércia é independente do tempo, o que torna as equações de movimento do corpo rígido substancialmente mais simples. A vida simplifica-se ainda mais com a escolha dos eixos fixos no corpo como eixos principais de inércia, o que será feito doravante. As componentes da Eq.(4.8.2) ao longo de eixos principais de inércia fixos no corpo são

$$\frac{dL_1}{dt} + \omega_2 L_3 - \omega_3 L_2 = N_1 \tag{4.8.3}$$

e as demais obtidas a partir desta por permutação cíclica dos subscritos. Fazendo uso de $L_1 = I_1\omega_1$, $L_2 = I_2\omega_2$, $L_3 = I_3\omega_3$ podemos escrever

$$\begin{aligned} I_1\dot{\omega}_1 - (I_2 - I_3)\omega_2\omega_3 &= N_1 , \\ I_2\dot{\omega}_2 - (I_3 - I_1)\omega_3\omega_1 &= N_2 , \\ I_3\dot{\omega}_3 - (I_1 - I_2)\omega_1\omega_2 &= N_3 . \end{aligned} \tag{4.8.4}$$

Estas são as célebres equações de movimento de Euler para um corpo rígido, e descrevem o movimento do eixo instantâneo de rotação (definido por $\boldsymbol{\omega}$) relativamente a um referencial fixo no corpo. Uma solução completa, que permita uma visualização do movimento, requer que se determine a orientação instantânea do corpo em relação a eixos inerciais externos. Mesmo no caso mais simples de rotação livre ($\mathbf{N} = 0$), a especificação completa do movimento de um corpo rígido arbitrário constitui um problema com elevado grau de dificuldade. O problema simplifica-se consideravelmente se o corpo tem um eixo de simetria. Em vez de tratar genericamente um corpo simétrico em rotação livre, consideraremos um exemplo especial que contém as principais características do caso geral.

■ **Exemplo 4.8.1.** Um prato de plástico ("frisbee") é arremessado quase horizontalmente para o ar, de tal modo que o prato gira ao mesmo tempo que seu plano oscila. Descrever o movimento de rotação do prato e mostrar que as oscilações são duas vezes mais rápidas do que a rotação do prato em torno do seu próprio eixo de simetria.

Solução. Seja I_3 o momento principal de inércia em torno do eixo de simetria do prato e $I_1 = I_2$ os momentos principais de inércia em torno de dois eixos mutuamente perpendiculares contidos no plano do prato e passando por seu centro de massa. Como é nulo o torque em relação ao centro de massa exercido pelo peso, as equações de Euler (4.8.4) reduzem-se a

$$\begin{aligned} I_1\dot{\omega}_1 - (I_1 - I_3)\omega_2\omega_3 &= 0 , \\ I_2\dot{\omega}_2 - (I_3 - I_1)\omega_3\omega_1 &= 0 , \\ I_3\dot{\omega}_3 &= 0 . \end{aligned} \tag{4.8.5}$$

Esta última equação mostra que ω_3 permanece invariável. Definindo a constante

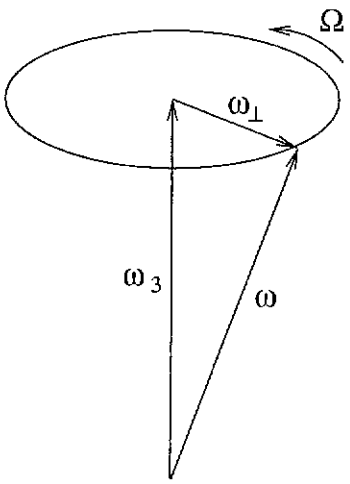


Figura 4.8: Precessão do eixo instantâneo de rotação de um corpo rígido simétrico.

$$\Omega = \frac{I_3 - I_1}{I_1}\omega_3 \ , \tag{4.8.6}$$

ficamos com

$$\dot{\omega}_1 = -\Omega\,\omega_2 \quad , \quad \dot{\omega}_2 = \Omega\,\omega_1 \ . \tag{4.8.7}$$

Destas equações deduz-se imediatamente

$$\ddot{\omega}_1 + \Omega^2\omega_1 = 0 \ . \tag{4.8.8}$$

Com uma escolha apropriada da origem do tempo, a solução desta última equação pode ser posta na forma

$$\omega_1 = A\cos\Omega t \ , \tag{4.8.9a}$$

donde

$$\omega_2 = A\,\mathrm{sen}\,\Omega t \ , \tag{4.8.9b}$$

sendo A uma constante arbitrária. O movimento executado pelo vetor ω está representado na Fig. 4.8. O vetor $\omega_\perp = \omega_1\hat{\xi}_1 + \omega_2\hat{\xi}_2$ gira com velocidade angular Ω em torno do eixo 3, ao mesmo tempo que ω executa um movimento de precessão em torno do eixo de simetria do corpo. O módulo de ω é constante e igual a $(\omega_3^2 + A^2)^{1/2}$, de modo que o eixo instantâneo de rotação descreve um cone à medida que precessa em torno do eixo de simetria. A precessão do vetor velocidade angular em torno do eixo de simetria aparece como uma oscilação do corpo. É importante sublinhar que a precessão de ω é relativa a eixos fixos no corpo, os quais, por sua vez, giram no espaço com velocidade angular ω, o que torna difícil a visualização do movimento.

A escolha das direções dos eixos principais de inércia no plano perpendicular ao eixo de simetria é arbitrária. Assim, podemos escolher num dado momento o eixo $x' \equiv x_1$ coincidente

com a linha nodal, de modo que no instante em questão $\psi = 0$ e as componentes da velocidade angular tornam-se mais simples:

$$\omega_1 = \dot{\theta} \quad , \quad \omega_2 = \dot{\phi}\,\mathrm{sen}\,\theta \quad , \quad \omega_3 = \dot{\phi}\cos\theta + \dot{\psi} \; . \tag{4.8.10}$$

Para simplificar ainda mais a análise, tomemos os eixo z do sistema de eixos fixos no espaço ao longo do vetor momento angular $\mathbf{L}$, que é uma constante de movimento. Uma vez que a linha nodal é perpendicular ao eixo z, temos (ver Fig. 3.4):

$$L_1 = 0 \quad , \quad L_2 = L\,\mathrm{sen}\,\theta \quad , \quad L_3 = L\cos\theta \; . \tag{4.8.11}$$

Assim, com a ajuda de (4.5.3), obtemos

$$\dot{\theta} = \frac{L_1}{I_1} = 0 \quad , \quad \dot{\phi} = \frac{\omega_2}{\mathrm{sen}\,\theta} = \frac{L}{I_1} \quad , \quad \omega_3 = \frac{L_3}{I_3} = \frac{L\cos\theta}{I_3} \; . \tag{4.8.12}$$

O ângulo de inclinação do eixo de simetria do corpo em relação ao vetor momento angular permanece constante. No caso do prato tem-se $I_3 = 2I_1$ em virtude do teorema do eixo perpendicular. Para um prato lançado quase horizontalmente o vetor momento angular é praticamente vertical, θ é pequeno e $\dot{\phi} \approx 2\omega_3$, isto é, a taxa de oscilação do prato (que coincide com a taxa de precessão $\dot{\phi}$) é duas vezes maior do que a taxa de rotação em torno do próprio eixo.[4] ■

Estabilidade do Movimento de Rotação Uniforme

Na ausência de torques, as equações de Euler admitem soluções tais que $\boldsymbol{\omega}$ possui componente diferente de zero e constante somente ao longo de um dos eixos principais de inércia. Por exemplo, $\omega_1 = \omega_1(0), \omega_2 = \omega_3 = 0$ é solução de (4.8.4) se $\mathbf{N} = 0$. Portanto, a rotação uniforme em torno de um dos eixos principais de inércia representa uma situação de equilíbrio para um corpo rígido livre (não sujeito a torques). Isto é importante para aplicações mecânicas, pois uma roda ou um volante só permanece girando uniformemente em torno de um dado eixo na ausência de torques se o eixo em questão for um eixo principal de inércia. Para manter um corpo girando com velocidade angular constante em torno de um eixo qualquer é necessário aplicar um torque ao corpo, resultando num maior desgaste no eixo. Uma questão interessante é se o movimento de rotação uniforme em torno de um eixo principal de inércia é um estado de equilíbrio estável ou instável.

■ **Exemplo 4.8.2.** Mostrar que só é estável a rotação uniforme em torno dos eixos principais de inércia de maior ou menor momento de inércia.

[4] Este problema desempenhou um papel bastante peculiar na trajetória científica do grande físico americano Richard Feynman. Em meados de 1947, Feynman era professor na Universidade Cornell. Depois de haver concluído o doutorado em Princeton e ter participado do Projeto Manhattan de construção da bomba atômica, em Los Alamos, ele atravessava uma fase improdutiva que o levou a pensar que estava acabado ("burned out") para a Física. Feynman preferia almoçar no restaurante estudantil porque gostava de olhar as moças bonitas. Um dia, durante o almoço, alguém de brincadeira arremessou para o ar um prato com o emblema do fundador da Universidade Cornell e Feynman observou como ele oscilava enquanto girava em torno do próprio eixo de simetria. Sem nada para fazer, ele formulou e resolveu as equações de movimento do prato oscilante-girante. Animado com a resolução do problema, Feynman dirigiu-se à sala de Hans Bethe e lhe contou o que tinha visto no restaurante e o que havia calculado. "Mas que importância tem isso?", perguntou Bethe. Feynman lhe disse que não tinha nenhuma importância nem ele ligava para isso, mas havia sido divertido. Ele decidiu que daquele momento em diante iria se se divertir com a Física e só faria o lhe desse prazer. Esse episódio reacendeu em Feynman a paixão lúdica pela Física e ajudou a pôr fim ao estado depressivo em que ele se encontrava (Feynman 1985; Mehra 1994).

Solução. Suponha que a rotação em torno de um eixo principal de inércia, digamos o terceiro eixo, seja ligeiramente perturbada de tal modo que apareçam componentes $\omega_1, \omega_2 \ll \omega_3$. O problema resume-se em descobrir se ω_1 e ω_2 permanecem pequenas, caso em que o equilíbrio será estável. As equações de Euler com $\mathbf{N} = 0$ escrevem-se, aproximadamente,

$$\begin{aligned} I_1\,\dot{\omega}_1 \;-\; (I_2 - I_3)\,\omega_2\,\omega_3 \;&=\; 0\ , \\ I_2\,\dot{\omega}_2 \;-\; (I_3 - I_1)\,\omega_3\,\omega_1 \;&=\; 0\ , \\ I_3\,\dot{\omega}_3 \;&=\; 0\ , \end{aligned} \tag{4.8.13}$$

onde o termo proporcional ao produto $\omega_1\omega_2$ foi desprezado por ser, por hipótese, muito pequeno. Conclui-se que ω_3 é aproximadamente constante e, diferenciando a primeira e usando a segunda das equações (4.8.13), deduz-se

$$\ddot{\omega}_1 + \Omega_0^2\omega_1 = 0\ , \tag{4.8.14}$$

onde

$$\Omega_0^2 = \frac{(I_3 - I_2)(I_3 - I_1)}{I_1 I_2}\omega_3^2\ . \tag{4.8.15}$$

Se $\Omega_0^2 > 0$, podemos escrever

$$\omega_1(t) = \omega_1(0)\cos\,\Omega_0 t \quad , \quad \omega_2(t) = \frac{\Omega_0 I_1}{(I_3 - I_2)\omega_3}\omega_1(0)\,\mathrm{sen}\,\Omega_0 t\ , \tag{4.8.16}$$

de modo que ω_1 e ω_2 oscilam harmonicamente em torno da situação de equilíbrio $\omega_1 = \omega_2 = 0$, a perturbação permanecendo pequena se, inicialmente, for suficientemente pequena. Mas $\Omega_0^2 > 0$ se: (i) $I_3 > I_2$ e $I_3 > I_1$; (ii) $I_3 < I_2$ e $I_3 < I_1$. Assim, a rotação uniforme é estável em torno dos eixos principais de inércia de maior ou menor momento de inércia. Se $\Omega_0^2 < 0$, ou seja, se $I_3 < I_2$ e $I_3 > I_1$, ou $I_3 > I_2$ e $I_3 < I_1$, ω_1 e ω_2 são funções exponenciais do tempo e a perturbação inicialmente pequena acaba por tornar-se grande: a rotação é instável se realizada em torno do eixo principal de inércia associado ao momento principal de inércia de valor intermediário. Este comportamento pode ser observado qualitativamente com qualquer objeto com a forma de um paralelepípedo: um bloco de madeira ou uma caixa de fósforos cheia. Basta segurar o objeto com o eixo escolhido na direção vertical, aplicar-lhe uma rotação em torno do referido eixo e deixá-lo cair livremente sob a ação do campo gravitacional. É fácil produzir uma rotação estável do corpo em torno dos eixos de menor ou maior momento de inércia. Mas as tentativas de pô-lo a girar em torno do eixo intermediário, por mais cuidadosas que sejam, degeneram rapidamente numa oscilação irregular, com o eixo de rotação perambulando aleatoriamente ao redor do corpo (Romer 1978). ■

A Construção de Poinsot

Quando não há torque, o vetor momento angular e a energia cinética são constantes de movimento:

$$\mathbf{L} = \overleftrightarrow{\mathbf{I}}\cdot\boldsymbol{\omega} = \textbf{constante}\ , \tag{4.8.17}$$

$$T = \frac{1}{2}\boldsymbol{\omega}\cdot\overleftrightarrow{\mathbf{I}}\cdot\boldsymbol{\omega} = \text{constante}\ . \tag{4.8.18}$$

Com base nestas constantes, e no caso geral em que o corpo rígido não é necessariamente simétrico, Poinsot concebeu uma elegante representação geométrica do movimento.

Definindo o vetor

$$\mathbf{r} = \frac{\boldsymbol{\omega}}{\sqrt{2T}} \tag{4.8.19}$$

a Eq.(4.8.18) toma a forma

$$\mathbf{r}\cdot \overleftrightarrow{\mathbf{I}} \cdot \mathbf{r} = 1 \ . \tag{4.8.20}$$

Por ser proporcional ao vetor velocidade angular, $\mathbf{r}$ é o vetor posição de um ponto do eixo instantâneo de rotação do corpo a contar da mesma origem O em relação à qual o tensor de inércia é calculado. A Eq.(4.8.20) define a superfície de um elipsóide porque, escrevendo-a relativamente a eixos principais de inércia do corpo rígido, ela reduz-se a

$$I_1 x_1^2 + I_2 x_2^2 + I_3 x_3^2 = 1 \ , \tag{4.8.21}$$

que representa um elipsóide com semi-eixos $I_1^{-1/2}$, $I_2^{-1/2}$, $I_3^{-1/2}$, conhecido como *elipsóide de inércia*. Num sistema cartesiano arbitrário, a Eq.(4.8.20) apenas representa o mesmo elipsóide com seus eixos inclinados em relação aos eixos coordenados. O vetor normal à superfície do elipsóide de inércia no ponto definido pela extremidade do vetor $\mathbf{r}$ é proporcional a

$$\begin{aligned}\boldsymbol{\nabla}(\mathbf{r}\cdot \overleftrightarrow{\mathbf{I}} \cdot \mathbf{r}) &= \Big(\hat{\mathbf{e}}_1 \frac{\partial}{\partial x_1} + \hat{\mathbf{e}}_2 \frac{\partial}{\partial x_2} + \hat{\mathbf{e}}_3 \frac{\partial}{\partial x_3}\Big)(I_1 x_1^2 + I_2 x_2^2 + I_3 x_3^2) \\ &= 2\,(I_1 x_1 \hat{\mathbf{e}}_1 + I_2 x_2 \hat{\mathbf{e}}_2 + I_3 x_3 \hat{\mathbf{e}}_3) = 2\, \overleftrightarrow{\mathbf{I}} \cdot \mathbf{r} = \sqrt{\frac{2}{T}}\,\mathbf{L} \ ,\end{aligned} \tag{4.8.22}$$

de modo que o plano tangente ao elipsóide de inércia no ponto $\mathbf{r}$ é perpendicular ao vetor $\mathbf{L}$, como ilustra a Fig. 4.9. A distância da origem do elipsóide de inércia ao plano tangente no ponto $\mathbf{r}$ é

$$d = \mathbf{r}\cdot\frac{\mathbf{L}}{L} = \frac{\boldsymbol{\omega}\cdot \overleftrightarrow{\mathbf{I}} \cdot \boldsymbol{\omega}}{L\sqrt{2T}} = \frac{\sqrt{2T}}{L} = \text{constante} \ . \tag{4.8.23}$$

O plano tangente ao elipsóide é completamente fixo no espaço, pois não apenas a direção normal ao plano é fixa como também a distância do plano à origem dos eixos coordenados é imutável, daí ser chamado de *plano invariável*. A posição do plano invariável relativamente ao centro do elipsóide de inércia é determinada pelas condições iniciais: $\mathbf{L}$ e T. Tendo em vista que o ponto de contato do elipsóide com o plano invariável pertence ao eixo instantâneo de rotação ($\mathbf{r}$ é paralelo a $\boldsymbol{\omega}$), o elipsóide de inércia rola sem deslizar sobre o plano invariável. À medida que o elipsóide rola sem deslizar, o ponto de contato descreve uma curva sobre o elipsóide, chamada de *polodia*,[5] e uma curva sobre o plano invariável,

[5]Do grego *pólos* (eixo) e *odós* (caminho).

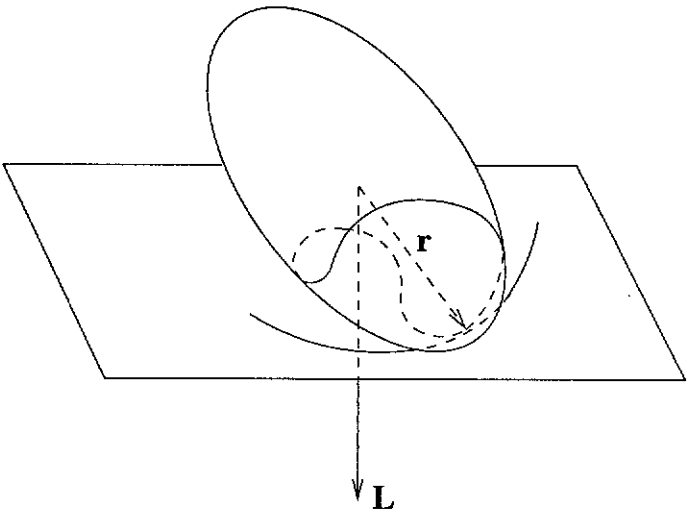


Figura 4.9: A construção de Poinsot.

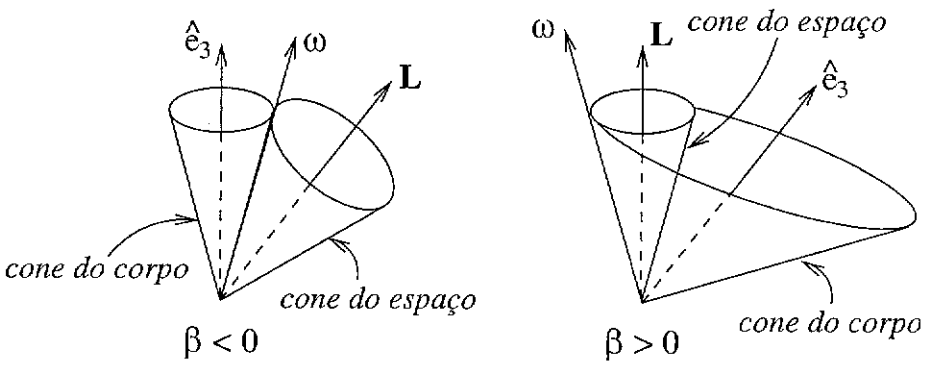


Figura 4.10: Os cones do corpo e do espaço.

denominada *herpolodia.*[6] Seria um tanto exagerado afirmar que essas palavras contribuem para embelezar a língua.

Quando o corpo é simétrico ($I_1 = I_2$), o elipsóide de inércia é um sólido de revolução em torno do eixo de simetria e constata-se sem dificuldade que a polodia e a herpolodia são circunferências. Com efeito, as Eqs.(4.8.5) mostram que ω_3 é constante e ω descreve um cone em torno do eixo de simetria chamado *cone do corpo.* De acordo com a Fig. 4.8, o semi-ângulo α_c do cone do corpo é dado por

$$\tan \alpha_c = \frac{\omega_\perp}{\omega_3} = \frac{A}{\omega_3} \ , \tag{4.8.24}$$

onde usamos as Eqs.(4.8.9). Por outro lado, o ângulo α_e entre $\boldsymbol{\omega}$ e $\mathbf{L}$ é constante pois

[6] O caminho serpentino.

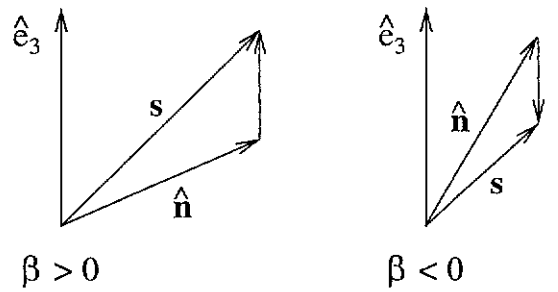


Figura 4.11: Vetores introduzidos para analisar as posições relativas dos cones do corpo e do espaço.

$$\cos\alpha_e = \frac{\boldsymbol{\omega}\cdot\mathbf{L}}{\omega L} = \frac{\boldsymbol{\omega}\cdot\overleftrightarrow{\mathbf{I}}\cdot\boldsymbol{\omega}}{\omega L} = \frac{2T}{\omega L} \ , \tag{4.8.25}$$

de modo que o eixo instantâneo de rotação traça um cone no espaço com semi-ângulo α_e. O vetor $\boldsymbol{\omega}$ pertence simultaneamente ao cone do corpo e ao cone do espaço, de sorte que o eixo instantâneo de rotação é a linha de contato entre os dois cones. Uma vez que os pontos do eixo instantâneo de rotação permanecem instantaneamente em repouso, para gerar o movimento o cone do corpo rola sem deslizar sobre o cone do espaço, que é fixo porque $\mathbf{L}$ é fixo no espaço. Assim, é claro que as curvas traçadas pelo eixo instantâneo de rotação sobre o elipsóide de inércia e sobre o plano invariável são circunferências. No caso geral de corpo assimétrico, as expressões analíticas da polodia e herpolodia são muito complicadas, envolvendo funções elípticas (Whittaker 1944).

À guisa de comentário final sobre o caso simétrico, notemos que o cone do corpo pode estar dentro ou fora do cone do espaço, conforme está ilustrado na Fig. 4.10. Seja $\boldsymbol{\omega} = \omega\hat{\mathbf{n}}$ e note que

$$\overleftrightarrow{\mathbf{I}} = I_1\hat{\mathbf{e}}_1\hat{\mathbf{e}}_1 + I_2\hat{\mathbf{e}}_2\hat{\mathbf{e}}_2 + I_3\hat{\mathbf{e}}_3\hat{\mathbf{e}}_3 = I_1\,\overleftrightarrow{\mathbf{1}} + (I_3 - I_1)\hat{\mathbf{e}}_3\hat{\mathbf{e}}_3 \ , \tag{4.8.26}$$

donde

$$\boldsymbol{s} \equiv \frac{\mathbf{L}}{\omega I_1} = \frac{\overleftrightarrow{\mathbf{I}}\cdot\boldsymbol{\omega}}{\omega I_1} = \frac{\boldsymbol{\omega}\,\overleftrightarrow{\mathbf{I}}\cdot\hat{\mathbf{n}}}{\omega I_1} = \hat{\mathbf{n}} + \beta\cos\alpha_c\,\hat{\mathbf{e}}_3 \ , \tag{4.8.27}$$

onde levamos em conta que $\cos\alpha_c = \hat{\mathbf{n}}\cdot\hat{\mathbf{e}}_3$ e definimos

$$\beta = \frac{I_3 - I_1}{I_1} \ . \tag{4.8.28}$$

Observando que $\boldsymbol{s}$ é paralelo a $\mathbf{L}$ e considerando a Fig. 4.11, conclui-se de imediato que o cone espacial está dentro do cone do corpo se $\beta > 0$ e fora do cone do corpo se $\beta < 0$. Este último caso corresponde a um corpo alongado ("acharutado") na direção do eixo de simetria, ao passo que $\beta > 0$ refere-se a um corpo achatado na direção do eixo de simetria.

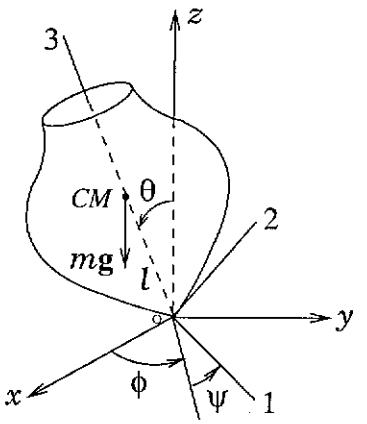


Figura 4.12: Pião simétrico com um ponto fixo.

4.9 Pião Simétrico com um Ponto Fixo

Genericamente, um pião simétrico é um corpo qualquer com $I_1 = I_2$. Suponhamos que ele seja capaz de girar em torno de um ponto fixo O do eixo de simetria, localizado a uma distância l do centro de massa, num campo gravitacional uniforme (Fig. 4.12). O movimento do pião será descrito em termos dos ângulos de Euler ϕ, θ, ψ. O ângulo θ mede a inclinação do eixo de simetria do pião relativamente à vertical, ao passo que ϕ descreve a precessão deste mesmo eixo em torno da direção vertical. Finalmente, ψ descreve a rotação do pião em torno do seu próprio eixo de simetria (vide a definição dos ângulos de Euler na Seção 3.3).

Lagrangiana e Constantes de Movimento

Em vez de empregar as equações de Euler (4.8.4), revela-se mais vantajoso lançar mão do formalismo lagrangiano para a análise do movimento do pião. Já que a única força aplicada é o peso, a lagrangiana escreve-se

$$L = T - V = \frac{1}{2} I_1 \left(\omega_1^2 + \omega_2^2\right) + \frac{1}{2} I_3 \, \omega_3^2 - mgl \cos\theta \ , \tag{4.9.1}$$

pois $I_1 = I_2$. Note que os eixos x', y', z' da Fig. 3.4 são os mesmos eixos $1, 2, 3$ da Fig. 4.12. Assim, recorrendo às Eqs.(3.4.30), a lagrangiana assume a forma

$$L = \frac{1}{2} I_1 \, \dot{\theta}^2 + \frac{1}{2} I_1 \, \dot{\phi}^2 \, \mathrm{sen}^2\theta + \frac{1}{2} I_3 \left(\dot{\psi} + \dot{\phi} \cos\theta\right)^2 - mgl \cos\theta \ . \tag{4.9.2}$$

As variáveis ψ e ϕ são cíclicas, logo os momentos conjugados correspondentes são as constantes de movimento p_ψ e p_ϕ:

$$\frac{\partial L}{\partial \dot{\psi}} = I_3 (\dot{\psi} + \dot{\phi} \cos\theta) = I_3 \, \omega_3 = p_\psi = \text{constante} \ , \tag{4.9.3}$$

$$\frac{\partial L}{\partial \dot{\phi}} = I_1\dot{\phi}\ \mathrm{sen}^2\theta + I_3\ \cos\theta(\dot{\psi}+\dot{\phi}\cos\theta) = I_1\dot{\phi}\ \mathrm{sen}^2\theta + p_\psi\cos\theta = p_\phi = \text{constante}\ . \quad (4.9.4)$$

■ **Exercício 4.9.1.** Mostre que o torque exercido pelo peso não tem componente ao longo da direção 3. Use, em seguida, as equações de Euler para deduzir a Eq.(4.9.3). ■

Uma vez que uma variação do ângulo ϕ implica uma rotação em torno do eixo vertical z, p_ϕ é a componente z do momento angular, L_z. A conservação de L_z, por seu turno, é facilmente compreendida notando que o torque em relação ao ponto fixo O devido ao peso não possui componente na direção vertical. Outra constante de movimento de importância fundamental é a energia:[7]

$$E = T + V = \frac{1}{2}I_1\,\dot{\theta}^2 + \frac{1}{2}I_1\,\dot{\phi}^2\ \mathrm{sen}^2\theta + \frac{1}{2}I_3\,(\dot{\psi}+\dot{\phi}\cos\theta)^2 + mgl\cos\theta\ . \quad (4.9.5)$$

Solução por Quadraturas

Essas três constantes de movimento permitem reduzir a solução do problema a quadraturas.[8] De fato, escrevendo

$$\dot{\psi} + \dot{\phi}\cos\theta = p_\psi/I_3\ , \quad (4.9.6)$$

$$\dot{\phi} = \frac{p_\phi - p_\psi\cos\theta}{I_1\,\mathrm{sen}^2\theta} \quad (4.9.7)$$

e introduzindo a nova constante

$$E' = E - \frac{p_\psi^2}{2I_3}\ , \quad (4.9.8)$$

deduzimos

$$E' = \frac{1}{2}I_1\dot{\theta}^2 + V_{\mathrm{ef}}(\theta)\ , \quad (4.9.9)$$

onde

$$V_{\mathrm{ef}}(\theta) = \frac{(p_\phi - p_\psi\cos\theta)^2}{2I_1\,\mathrm{sen}^2\theta} + mgl\cos\theta\ . \quad (4.9.10)$$

Separando variáveis na Eq.(4.9.9) obtém-se

$$\int_{\theta_0}^{\theta}\frac{d\theta}{[E' - V_{\mathrm{ef}}(\theta)]^{1/2}} = \sqrt{\frac{2}{I_1}}\,t\ , \quad (4.9.11)$$

[7] A lagrangiana não depende explicitamente do tempo e a energia cinética é função homogênea do segundo grau das velocidades, de modo que a integral de Jacobi h é a energia (vide Seção 2.6).

[8] Resolver o problema por quadraturas significa exprimir a solução do problema em termos de integrais de funções conhecidas.

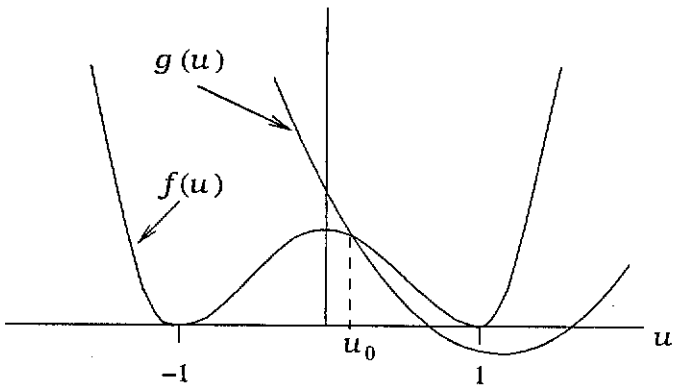


Figura 4.13: A interseção das curvas definidas por $f(u)$ e $g(u)$ fornece o único valor de $\theta_0 = \cos^{-1} u_0$ que anula $V'_{\rm ef}(\theta)$ no intervalo fisicamente aceitável. O caso representado corresponde a $p_\psi p_\phi > 0$.

onde θ_0 é o valor de θ no instante inicial $t = 0$.[9] Em princípio, portanto, uma única quadratura (integração) permite determinar t como função de θ, uma inversão adicional fornecendo $\theta(t)$. Uma vez encontrada a dependência temporal de θ, determina-se $\phi(t)$ por uma integração da Eq.(4.9.7). Finalmente, dispondo de $\theta(t)$ e $\phi(t)$, obtém-se $\psi(t)$ por uma última integração utilizando a Eq.(4.9.6). A integral em (4.9.11), no entanto, não pode ser expressa em termos de funções elementares – salvo para condições iniciais muito particulares –, de modo que, em geral, θ exprime-se em função de t em termos de funções elípticas de Weierstrass ou de Jacobi.[10] Afortunadamente, as características principais do movimento do pião podem ser inferidas sem a necessidade de lidar com essas funções relativamente complicadas.

Potencial Efetivo

O comportamento de $\theta(t)$ pode ser analisado pelo método do potencial efetivo com a ajuda da equação (4.9.9). Consideremos uma situação geral em que $|p_\psi| \neq |p_\phi|$ (um exemplo do caso especial $p_\psi = p_\phi$ será tratado mais adiante). O intervalo fisicamente aceitável para θ é $[0, \pi]$, e de (4.9.10) deduz-se que $V_{\rm ef}(\theta) \longrightarrow \infty$ para $\theta \longrightarrow 0$ ou π. Por outro lado, a derivada do potencial efetivo anula-se para θ_0 tal que

$$V'_{\rm ef}(\theta_0) = -mgl \operatorname{sen}\theta_0 + \frac{(p_\phi - p_\psi \cos\theta_0)(p_\psi - p_\phi \cos\theta_0)}{I_1 \operatorname{sen}^3\theta_0} = 0 \ . \qquad (4.9.12)$$

Fazendo $u = \cos\theta_0$, esta última equação equivale a

$$mglI_1(1-u^2)^2 = (p_\phi - p_\psi u)(p_\psi - p_\phi u) \ . \qquad (4.9.13)$$

[9] Se $\dot\theta < 0$ no instante $t = 0$ deve-se tomar a raiz quadrada negativa no primeiro membro de (4.9.11).

[10] Uma excelente introdução às funções elípticas de Jacobi encontra-se em Synge & Griffith (1959). Para outros detalhes a respeito de integrais elípticas e funções elípticas, Spiegel (1971) é uma referência que prima pela concisão e objetividade.

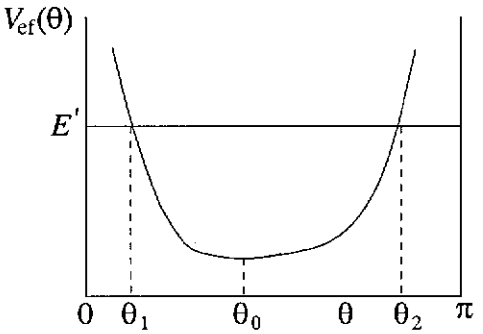


Figura 4.14: O potencial efetivo $V_{\rm ef}(\theta)$.

Denotando o primeiro membro de (4.9.13) por $f(u)$ e o segundo por $g(u)$, podemos determinar graficamente o número de interseções das curvas definidas por $f(u)$ e $g(u)$ para $-1 \leq u \leq 1$. As raízes de $f(u)$ são ± 1, a tangente ao gráfico de $f(u)$ é horizontal em $u = 0, \pm 1$ e $f(u) \longrightarrow \infty$ para $u \longrightarrow \pm\infty$. Por outro lado, as raízes de $g(u)$ são $u_1 = p_\phi/p_\psi$ e $1/u_1$, de modo que se u_1 pertence ao intervalo $(-1, 1)$ então $1/u_1$ não pertence a $(-1, 1)$ e vice-versa. Em outras palavras, há apenas uma raiz de $g(u)$ no intervalo fisicamente significativo. Estas informações bastam para traçar a Fig. 4.13, que torna claro que a equação (4.9.12) possui uma única solução u_0 no intervalo físico, ou, equivalentemente, há um único valor $\theta_0 \in [0, \pi]$ que anula $V'_{\rm ef}(\theta)$. Inferimos, assim, que a forma de $V_{\rm ef}(\theta)$ é aquela esboçada na Fig. 4.14.

Características Gerais do Movimento do Pião

Em geral, dado um valor de E' o ângulo θ oscila periodicamente entre os valores θ_1 e θ_2 indicados na Fig. 4.14. Diz-se que o eixo de simetria do pião executa uma *nutação* em relação à vertical. Durante a nutação a velocidade angular de precessão do eixo do pião em torno da direção vertical é, em cada instante,

$$\dot{\phi} = \frac{p_\phi - p_\psi \cos\theta}{I_1 \,\mathrm{sen}^2\theta} = p_\psi \frac{\frac{p_\phi}{p_\psi} - \cos\theta}{I_1 \,\mathrm{sen}^2\theta} \ . \tag{4.9.14}$$

Há diversas possibilidades, a saber:

(i) $|p_\phi| > |p_\psi|$. A taxa de precessão $\dot{\phi}$ tem sempre o mesmo sinal durante todo o tempo.

(ii) $|p_\phi| < |p_\psi|$. Introduzindo o ângulo auxiliar $\theta' \in [0, \pi]$ definido por $\cos\theta' = \frac{p_\phi}{p_\psi}$, temos $\dot{\phi} = p_\psi(\cos\theta' - \cos\theta)/I_1 \,\mathrm{sen}^2\theta$. Para determinar a posição de θ', note que a equação (4.9.12) fornece $V'_{\rm ef}(\theta') = -mgl \,\mathrm{sen}\,\theta' < 0$, de modo que $\theta' < \theta_0$. Se $\theta' < \theta_1$ então $\dot{\phi}$ tem sempre o mesmo sinal durante todo o movimento. Neste caso, assim como no anterior, o eixo de simetria do pião traça, sobre uma esfera centrada no ponto O, uma curva que assemelha-se à mostrada na Fig. 4.15(a). Se $\theta' > \theta_1$, $\dot{\phi}$ troca de sinal durante a nutação e o movimento corresponde ao indicado na Fig. 4.15(b). A situação retratada na Fig. 4.15(c) corresponde a $\theta' = \theta_1$. Este último caso não é tão excepcional quanto parece.

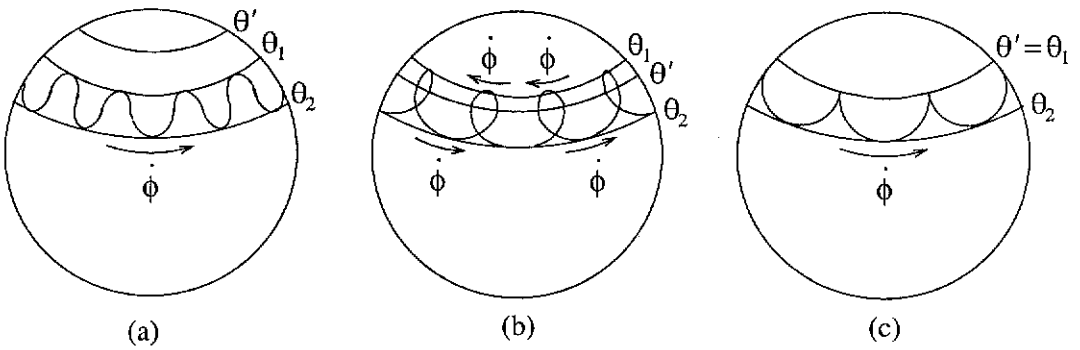


Figura 4.15: As combinações possíveis dos movimentos de precessão e nutação.

Suponha, por exemplo, que as condições iniciais sejam $\dot{\theta} = 0, \dot{\phi} = 0, \dot{\psi} = \omega_3$, de modo que o único movimento inicial do pião consiste numa rotação em torno de seu próprio eixo de simetria. Com essas condições iniciais tem-se

$$p_\psi = I_3\omega_3 \quad , \quad p_\phi = p_\psi \cos\theta_1 \ ,$$

donde

$$\cos\theta' = \frac{p_\phi}{p_\psi} = \cos\theta_1 \implies \theta' = \theta_1 \ .$$

Se $E' = V_{\text{ef}}(\theta_0)$ o ângulo θ permanece constante e igual a θ_0, não ocorre nutação e o eixo de simetria do pião executa uma *precessão regular* em torno da vertical com a velocidade angular constante

$$\dot{\phi} = \frac{p_\phi - p_\psi \cos\theta_0}{I_1 \,\text{sen}^2\theta_0} \ . \tag{4.9.15}$$

Há, em geral, dois valores possíveis para $\dot{\phi}$ porque a equação (4.9.12) é quadrática em

$$\sigma = p_\phi - p_\psi \cos\theta_0 \ . \tag{4.9.16}$$

De fato,

$$p_\psi - p_\phi \cos\theta_0 = p_\psi \,\text{sen}^2\theta_0 - \sigma\cos\theta_0 \ , \tag{4.9.17}$$

de modo que (4.9.12) torna-se

$$\sigma^2\cos\theta_0 - \sigma p_\psi \,\text{sen}^2\theta_0 + mglI_1 \,\text{sen}^4\theta_0 = 0 \ , \tag{4.9.18}$$

cujas soluções são

$$\sigma = \frac{I_3\omega_3 \,\text{sen}^2\theta_0}{2\cos\theta_0}\left[1 \pm \sqrt{1 - \frac{4mglI_1\cos\theta_0}{I_3^2\omega_3^2}}\right] \ , \tag{4.9.19}$$

onde usamos $p_\psi = I_3\omega_3$. Se $\theta_0 < \pi/2$, de modo que $\cos\theta_0 > 0$, a velocidade angular do pião em torno de seu eixo de simetria não pode ser inferior a um valor mínimo[11] para que o pião possa precessar regularmente com ângulo de inclinação θ_0:

$$\omega_{\min} = \left(\frac{4mglI_1\cos\theta_0}{I_3^2}\right)^{1/2} . \tag{4.9.20}$$

Se $\omega_3 > \omega_{\min}$ há duas soluções para $\sigma = p_\phi - p_\psi\cos\theta_0$, correspondendo a uma precessão lenta e outra rápida. Para $\omega_3 \gg \omega_{\min}$ temos os seguintes valores aproximados:

$$\dot\phi_{\text{rápida}} \approx \frac{I_3\omega_3}{I_1\cos\theta_0} \quad , \quad \dot\phi_{\text{lenta}} \approx \frac{mgl}{I_3\omega_3} . \tag{4.9.21}$$

■ **Exercício 4.9.2.** Deduza estes dois últimos resultados aproximados aplicando, se necessário, uma expansão binomial à Eq.(4.9.19). ■

Estabilidade de um Pião Vertical

O último problema importante que investigaremos é o da estabilidade de um pião posto a girar na posição vertical, também conhecido como pião dormente.[12] As condições iniciais $\theta(0) = 0, \dot\psi(0) = \omega_3$ implicam, pela equação (4.9.4), $p_\phi = p_\psi$. Em tais circunstâncias, o potencial efetivo torna-se

$$V_{\text{ef}}(\theta) = \frac{p_\psi^2}{2I_1}\frac{(1-\cos\theta)^2}{\text{sen}^2\theta} + mgl\cos\theta = \frac{p_\psi^2}{2I_1}\tan^2\frac{\theta}{2} + mgl\cos\theta . \tag{4.9.22}$$

Para pequenas nutações relativamente à posição vertical,

$$V_{\text{ef}}(\theta) \approx \frac{p_\psi^2}{2I_1}\left(\frac{\theta}{2}\right)^2 + mgl\left(1-\frac{\theta^2}{2}\right) = mgl + \frac{1}{2}k\theta^2 , \tag{4.9.23}$$

onde

$$k = \frac{I_3^2\omega_3^2}{4I_1} - mgl . \tag{4.9.24}$$

Para pequenos valores de θ o gráfico de $V_{\text{ef}}(\theta)$ é uma parábola com um mínimo em $\theta = 0$, se $k > 0$, e um máximo em $\theta = 0$, se $k < 0$. Conseqüentemente, o movimento do pião com eixo de simetria vertical será estável se $k > 0$, ou seja, se

$$\omega_3 \geq \omega_c = \left(\frac{4mglI_1}{I_3^2}\right)^{1/2} . \tag{4.9.25}$$

[11] No caso de um giroscópio suspenso pelo topo, $\theta_0 > \pi/2$ e a precessão regular é possível qualquer que seja o valor de ω_3.

[12] O leitor é remetido a Whittaker (1944) para outros aspectos do movimento do pião, bem como para a solução exata em termos de funções elípticas.

Em outras palavras, a velocidade angular do pião em torno do seu próprio eixo possui um valor crítico ω_c abaixo do qual o movimento de rotação do pião na posição vertical torna-se instável. Se $\omega_3 > \omega_c$ as pequenas oscilações do eixo em torno da direção vertical permanecem pequenas, o eixo de simetria do pião continua praticamente vertical e ele rodopia serenamente, parecendo parado, dormente. Na prática, se um pião é posto a girar rapidamente e quase verticalmente, ele fica "adormecido" até que, por causa do atrito, ω_3 torna-se menor do que ω_c, quando o pião "desperta" e começa a cambalear, com seu eixo de simetria afastando-se cada vez mais da vertical, até cair inapelavelmente. Qualquer pessoa que tenha brincado com piões certamente já terá observado esse comportamento.

PROBLEMAS

4.1. O cubo do Exemplo 4.6.1 é posto a girar em torno da aresta que coincide com o eixo z. Determine o vetor momento angular do cubo e o ângulo que ele faz com o vetor velocidade angular.

4.2. Defina o produto vetorial pela direita do vetor $\mathbf{A}$ por uma díada e, por linearidade, usando (4.2.5) ou (4.2.6), estenda a definição a um diádico genérico $\overleftrightarrow{\mathbf{T}}$. Mostre que vale a fórmula $(\overleftrightarrow{\mathbf{T}} \times \mathbf{A}) \cdot \mathbf{B} = \overleftrightarrow{\mathbf{T}} \cdot (\mathbf{A} \times \mathbf{B})$ para um vetor arbitrário $\mathbf{B}$. Esta fórmula pode ser utilizada para definir $\overleftrightarrow{\mathbf{T}} \times \mathbf{A}$ por sua aplicação a um vetor $\mathbf{B}$ qualquer.

4.3. Os momentos principais de inércia I_1, I_2, I_3 de um corpo rígido em relação ao centro de massa são todos distintos. Se o centro de rotação do corpo passar a ser um ponto O deslocado de um vetor $\mathbf{s}$ em relação ao centro de massa, mostre que apenas se $\mathbf{s}$ for ao longo de um dos eixos $1, 2, 3$ é que os eixos principais de inércia passando pelo ponto O serão paralelos aos eixos $1, 2, 3$.

4.4. Mostre que cada momento de inércia nunca é maior do que a soma dos outros dois.

4.5. A aceleração centrífuga $\mathbf{a}_c = \omega \times (\mathbf{r} \times \omega)$ depende linearmente de $\mathbf{r}$, logo pode ser escrita na forma $\mathbf{a}_c = \overleftrightarrow{\mathbf{T}}_c \cdot \mathbf{r}$. Exprima o tensor $\overleftrightarrow{\mathbf{T}}_c$ em termos de ω e obtenha a matriz das componentes de $\overleftrightarrow{\mathbf{T}}_c$ num sistema cartesiano qualquer.

4.6. Determine os eixos e momentos principais de inércia em relação ao vértice de um cone homogêneo de altura h e raio da base R. Sabendo que o centro de massa do cone encontra-se a uma distância $3h/4$ do vértice, obtenha os eixos e momentos principais de inércia em relação ao centro de massa.

4.7. (i) Dado um tensor arbitrário $\overleftrightarrow{\mathbf{T}}$, prove que

$$\left(\frac{d\overleftrightarrow{\mathbf{T}}}{dt}\right)_{inercial} = \left(\frac{d\overleftrightarrow{\mathbf{T}}}{dt}\right)_{corpo} + \omega \times \overleftrightarrow{\mathbf{T}} - \overleftrightarrow{\mathbf{T}} \times \omega \ .$$

(ii) Usando $\mathbf{L} = \overleftrightarrow{\mathbf{I}} \cdot \omega$, $(d\mathbf{L}/dt)_{inercial} = \mathbf{N}$ e o resultado do item anterior, deduza as equações de Euler para o movimento de um corpo rígido.

4.8. Uma nave espacial simétrica move-se no espaço sideral. Motores simetricamente situados aplicam um torque constante N_3 ao longo do eixo de simetria. (i) Supondo que $\omega_3(0) = 0$, determine $\omega_3(t)$. (ii) Prove que $\omega_1^2 + \omega_2^2$ é constante de movimento. (iii) Tomando como condições iniciais $\omega_1(0) = 0$ e $\omega_2(0) = \Omega$, determine $\omega_1(t)$ e $\omega_2(t)$. Descreva o movimento executado pelo vetor velocidade angular relativamente aos eixos principais de inércia.

4.9. Mostre que energia cinética em relação ao centro de massa ou a um ponto fixo O satisfaz a equação $dT/dt = \mathbf{N} \cdot \omega$, conhecida como *teorema da energia cinética.*

4.10. Uma haste homogênea de comprimento $2l$ e massa m está presa por uma extremidade a uma mola cuja constante elástica é k, e que pende de um um suporte fixo. A haste pode oscilar num plano vertical e a mola está restrita a mover-se apenas para cima ou para baixo. Escreva a lagrangiana e as equações de movimento desse sistema.

4.11 Uma porta homogênea de massa m, largura l, altura h e espessura desprezível está se fechando com velocidade angular constante ω (ver figura). (i) Escolhendo o eixo

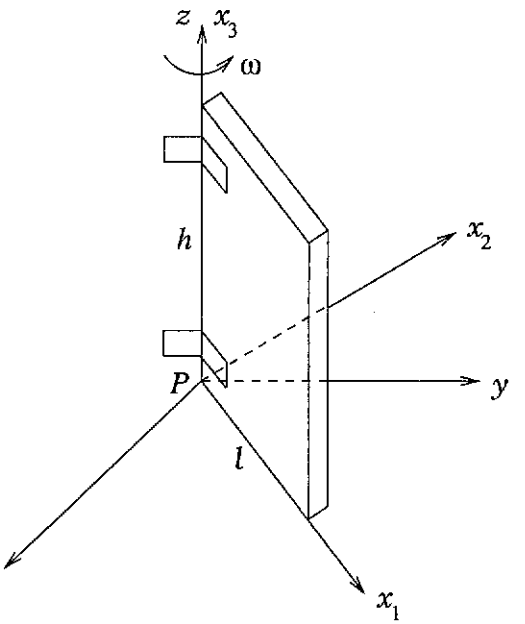


Figura 4.16: Problema 4.11.

z coincidente com o eixo de rotação com origem no ponto P, encontre as componentes do vetor momento angular ao longo dos eixos x_1, x_2, x_3 fixos na porta. Sugestão: Determine as componentes do tensor de inércia relativas ao ponto P vistas do sistema cartesiano atado à porta. (ii) Determine o vetor torque em relação ao ponto P. Sugestão: transforme as componentes de $\mathbf{L}$ para eixos x, y, z fixos no espaço e use $\mathbf{N} = d\mathbf{L}/dt$.

4.12 Um automóvel parte do repouso com aceleração constante a e uma das portas totalmente aberta (fazendo 90^o com o corpo do veículo). Se a porta é um retângulo homogêneo de largura l e altura h, prove que o tempo transcorrido até o fechamento da porta é

$$t = \sqrt{\frac{l}{3a}} \int_0^{\pi/2} \frac{d\theta}{\sqrt{\operatorname{sen}\theta}} \ .$$

Se $l = 1,2m$ e $a = 0,3m/s^2$, mostre que $t \approx 3s$. Sugestões: (i) No referencial do carro, o problema é formalmente idêntico ao de uma porta girando em torno de um eixo fixo horizontal sob a ação de um campo gravitacional constante. (ii) Calcule numericamente a integral ou procure o seu valor em alguma tabela.

4.13. Um pião simétrico com um ponto fixo entra em movimento da seguinte maneira:

$$\theta = 60^o \quad , \quad \dot{\theta} = 0 \quad , \quad \dot{\phi} = 2\left(\frac{mgl}{3I_1}\right)^{1/2} \quad , \quad \dot{\psi} = (3I_1 - I_3)\left(\frac{mgl}{3I_1 I_3^2}\right)^{1/2} \ .$$

(i) Encontre os valores de p_ϕ e p_ψ, e determine $V_{\text{ef}}(\theta)$. Esboce um gráfico de $V_{\text{ef}}(\theta)$

para $0 \leq \theta \leq \pi$ e faça uma análise qualitativa do movimento de nutação. (ii) Mostre que a equação (4.9.9) para θ pode ser posta na forma

$$\dot{u}^2 = \frac{mgl}{I_1}(1-u)^2(2u-1) \ ,$$

onde $u = \cos\theta$. Prove que a solução para $\theta(t)$ é

$$\sec\theta = 1 + \operatorname{sech}\left[\left(\frac{mgl}{I_1}\right)^{1/2} t\right] \ .$$

O comportamento de θ como função de t está de acordo com a análise qualitativa anterior? Estude, também, o comportamento de $\dot{\phi}$ e $\dot{\psi}$.

Capítulo 5

PEQUENAS OSCILAÇÕES

> Time travels in divers paces with divers persons. I'll tell you who Time ambles withal, who Time trots withal, who Time gallops withal, and who he stands still withal.
>
> William Shakespeare, *As You Like It*

A teoria das pequenas oscilações em torno de posições de equilíbrio estável encontra aplicações na espectroscopia molecular, no estudo de vibrações mecânicas, na acústica e na teoria dos circuitos elétricos, entre outras áreas. O conceito de *modo normal de vibração*, introduzido na teoria, é de enorme importância em praticamente todos os ramos da Física, da teoria quântica de campos à física da matéria condensada. Se os desvios da posição de equilíbrio forem suficientemente pequenos, em geral o movimento poderá ser descrito como o de uma coleção de osciladores harmônicos acoplados.

5.1 Caso Unidimensional

Consideremos um sistema conservativo com um único grau de liberdade descrito pela coordenada generalizada q. Sem perda significativa de generalidade, suponhamos que a relação entre as coordenadas cartesianas e a coordenada generalizada q não envolva o tempo, de modo que a lagrangiana tem a forma[1]

$$L = \frac{1}{2}\alpha(q)\,\dot{q}^2 \; - \; V(q) \;\; , \tag{5.1.1}$$

com a função $\alpha(q)$ positiva.

■ **Exemplo 5.1.1.** Uma conta de massa m desliza sem atrito ao longo de um fio rígido contido num plano vertical e na forma de uma parábola com equação $y = x^2/R$. O sistema encontra-se num campo gravitacional constante $\mathbf{g} = -g\hat{\mathbf{j}}$. Usando $q = x$ como coordenada generalizada, temos

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - mgy = \frac{m}{2}\left(1 + \frac{4x^2}{R^2}\right)\dot{x}^2 - \frac{mg}{R}x^2 \;\; , \tag{5.1.2}$$

[1] Considere a versão unidimensional das equações (2.6.4) e (2.6.5).

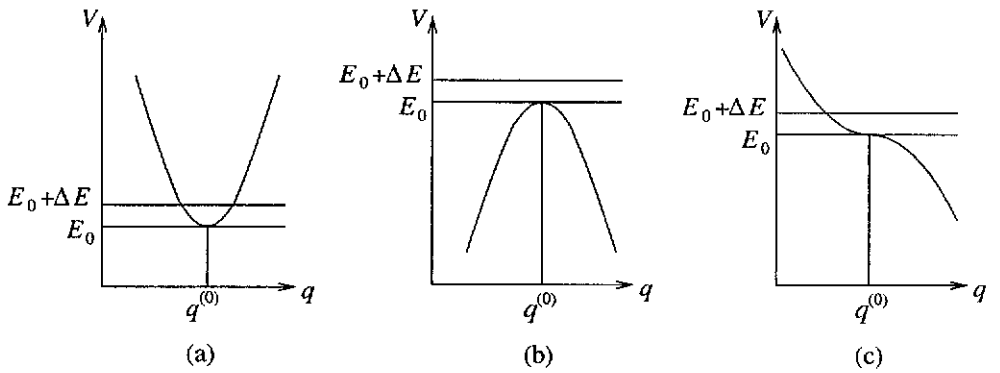


Figura 5.1: Em (a) o equilíbrio é estável, mas em (b) ou (c) é instável.

que é da forma (5.1.1) com $\alpha(x) = m(1 + 4x^2/R^2)$. ■

Configurações de Equilíbrio

Diz-se que a configuração $q = q^{(0)}$ é um ponto de equilíbrio se

$$\left(\frac{dV}{dq}\right)_{q=q^{(0)}} = 0 \ . \tag{5.1.3}$$

De fato, a equação de movimento produzida pela lagrangiana (5.1.1) é

$$\alpha(q)\ddot{q} + \frac{\alpha'(q)}{2}\dot{q}^2 + \frac{dV}{dq} = 0 \ , \tag{5.1.4}$$

que possui a solução $q = q^{(0)} =$ constante desde que (5.1.3) seja verificada. Em outras palavras, se num dado instante o sistema encontra-se na configuração $q = q^{(0)}$ com velocidade $\dot{q} = 0$, o sistema permanecerá nessa situação para sempre. Por exemplo, um pêndulo inicialmente em repouso na posição de equilíbrio ($\theta = 0$) assim perdurará indefinidamente.

Critério de Estabilidade

O equilíbrio é dito *estável* se uma perturbação suficientemente pequena não é capaz de afastar grandemente o sistema da condição de equilíbrio, mas provoca pequenas oscilações em torno da configuração de equilíbrio. Isto ocorre se o potencial passa por um mínimo na posição de equilíbrio. Se V passa por um máximo ou ponto de inflexão, o equilíbrio é instável. Como se depreende da Fig. 5.1, se uma pequena velocidade for imprimida à partícula na configuração de equilíbrio, aumentando ligeiramente sua energia, o deslocamento permanecerá pequeno se o potencial V for mínimo, mas não será pequeno, podendo crescer indefinidamente, quando V é máximo ou passa por um ponto de inflexão. O critério de estabilidade é

$$\left(\frac{d^2V}{dq^2}\right)_{q=q^{(0)}} > 0 \ , \tag{5.1.5}$$

que garante que V passa por um mínimo local em $q = q^{(0)}$.

Pequenas Oscilações em Torno do Equilíbrio Estável

Estaremos interessados em pequenas vibrações em torno de um ponto de equilíbrio estável. Defina

$$q = q^{(0)} + \eta \tag{5.1.6}$$

onde η é o deslocamento da posição de equilíbrio. Supondo η suficientemente pequeno, podemos aproximar $V(q)$ nas vizinhanças da posição de equilíbrio por

$$V(q) = V(q^{(0)} + \eta) = V(q^{(0)}) + \left(\frac{dV}{dq}\right)_{q=q^{(0)}} \eta + \frac{1}{2}\left(\frac{d^2V}{dq^2}\right)_{q=q^{(0)}} \eta^2 + \cdots \ . \tag{5.1.7}$$

O termo linear em η nesta última equação é nulo em virtude da Eq.(5.1.3). Portanto, até segunda ordem em η,

$$V(q) = V_0 + \frac{1}{2}k^{(0)}\eta^2 \ , \tag{5.1.8}$$

onde

$$k^{(0)} = \left(\frac{d^2V}{dq^2}\right)_{q=q^{(0)}} > 0 \tag{5.1.9}$$

e $V_0 = V(q^{(0)})$ é constante e igual ao valor da energia potencial na configuração de equilíbrio.

Coerentemente, façamos uma expansão da energia cinética até segunda ordem em η e $\dot{\eta}$. Temos

$$\alpha(q) = \alpha(q^{(0)} + \eta) = \alpha(q^{(0)}) + \left(\frac{d\alpha}{dq}\right)_{q=q^{(0)}} \eta + \cdots \ . \tag{5.1.10}$$

Porém, como $\dot{q} = \dot{\eta}$, vemos que o termo linear em η nesta última equação daria lugar na energia cinética a um termo proporcional a $\eta\dot{\eta}^2$, que é cúbico nas quantidades pequenas η e $\dot{\eta}$. Portanto, mantendo no máximo termos quadráticos em η e $\dot{\eta}$, a lagrangiana (5.1.1) escreve-se

$$L = \frac{\alpha^{(0)}}{2}\dot{\eta}^2 - \frac{k^{(0)}}{2}\eta^2 \ , \tag{5.1.11}$$

com

$$\alpha^{(0)} = \alpha(q^{(0)}) \ , \tag{5.1.12}$$

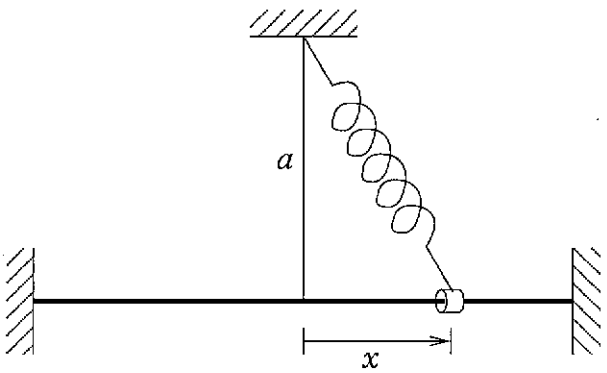


Figura 5.2: Sistema mecânico do Exemplo 5.1.2.

tendo sido descartada a constante V_0, pois não afeta a equação de movimento para η. Constata-se imediatamente que (5.1.11) tem a forma da lagrangiana de um oscilador harmônico de "massa" $\alpha^{(0)}$ e "constante elástica" $k^{(0)}$, de modo que η executa oscilações harmônicas em torno de $\eta = 0$ com freqüência angular

$$\omega = \left(\frac{k^{(0)}}{\alpha^{(0)}}\right)^{1/2} . \tag{5.1.13}$$

No Exemplo 5.1.1 claramente $x = 0$ é ponto de equilíbrio estável, pois $V'(0) = (2mgx/R)_{x=0} = 0$ e $V''(0) = 2mg/R > 0$. Portanto, de acordo com (5.1.13), a freqüência das pequenas oscilações em torno de $x = 0$ será $\omega = (2g/R)^{1/2}$.

É necessário sublinhar a hipótese de que a expansão do potencial em torno da posição de equilíbrio seja dominada pelo termo de segunda ordem. Nos casos anômalos em que o termo de segunda ordem está ausente, o movimento oscilatório em torno das posições de equilíbrio estável será periódico mas não será harmônico simples, de modo que a análise acima não se aplica.

■ **Exemplo 5.1.2.** Uma conta de massa m desliza ao longo de uma haste cilíndrica lisa ligada a uma mola com constante de força k e comprimento natural l, conforme a Fig. 5.2. Determinar as posições de equilíbrio e classificá-las quanto à estabilidade nos casos $l > a$ e $l < a$.

Solução. Usando como coordenada generalizada o deslocamento horizontal x, a lagrangiana escreve-se

$$L = \frac{m}{2}\dot{x}^2 - \frac{k}{2}(\sqrt{a^2 + x^2} - l)^2 . \tag{5.1.14}$$

As posições de equilíbrio satisfazem

$$\frac{dV}{dx} = k(\sqrt{a^2 + x^2} - l)\frac{x}{\sqrt{a^2 + x^2}} = 0 , \tag{5.1.15}$$

cujas soluções são $x = 0$ e $x = \pm\sqrt{l^2 - a^2}$ (se $l > a$). Para investigar a estabilidade devemos determinar o sinal de

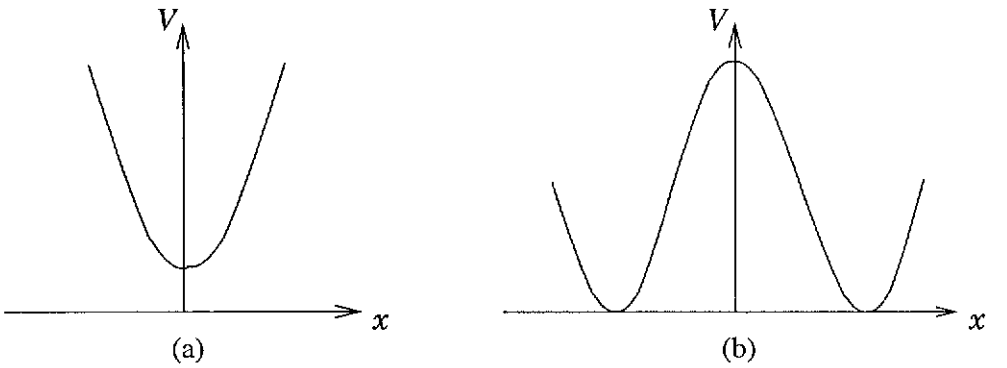


Figura 5.3: Potencial $V(x) = (k/2)(\sqrt{a^2+x^2} - l)^2$ nos casos (a) $a > l$ e (b) $a < l$.

$$\frac{d^2V}{dx^2} = k\frac{x^2}{a^2+x^2} + k\Big(\sqrt{a^2+x^2} - l\Big)\frac{a^2}{(a^2+x^2)^{3/2}} \tag{5.1.16}$$

nos pontos de equilíbrio. (i) Caso $a > l$. Somente $x = 0$ é ponto de equilíbrio, o qual é estável, pois

$$\left(\frac{d^2V}{dx^2}\right)_{x=0} = k\Big(1 - \frac{l}{a}\Big) > 0 \ , \tag{5.1.17}$$

e a freqüência das pequenas oscilações em torno de $x = 0$ é

$$\omega = \sqrt{\frac{k}{m}(1 - \frac{l}{a})} \ . \tag{5.1.18}$$

(ii) Caso $a < l$. Há três posições de equilíbrio, $x = 0$ ou $x = \pm\sqrt{l^2 - a^2}$. O ponto $x = 0$ agora é posição de equilíbrio instável porque $V''(0) < 0$. Por outro lado, as posições $x = \pm\sqrt{l^2 - a^2}$ são de equilíbrio estável porque

$$\left(\frac{d^2V}{dx^2}\right)_{x=\pm\sqrt{l^2-a^2}} = k(1 - \frac{a^2}{l^2}) > 0 \ , \tag{5.1.19}$$

sendo

$$\omega = \sqrt{\frac{k}{m}\Big(1 - \frac{a^2}{l^2}\Big)} \tag{5.1.20}$$

a freqüência das pequenas oscilações em torno dessas posições. Note que se $a = l$ nos deparamos com a situação anômala mencionada anteriormente: a única posição de equilíbrio é $x = 0$ mas o termo dominante na expansão do potencial em torno de $x = 0$ é de quarta ordem. ■

Quebra Espontânea de Simetria

Os gráficos do potencial $V(x)$ para $a > l$ e $a < l$ estão mostrados na Fig. 5.3. Este sistema é muito interessante porque exibe o fenômeno de *quebra espontânea de simetria.*

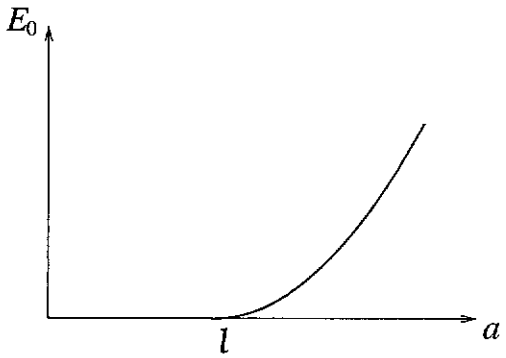


Figura 5.4: Energia do estado fundamental do Exemplo 5.1.2 em função da distância a.

Note que a lagrangiana (5.1.14) é invariante sob a transformação $x \to -x$. Se $a > l$ a solução de mínima energia é $x = 0$, com esta única configuração de equilíbrio estável exibindo a mesma propriedade de invariância que a lagrangiana. No entanto, para $a < l$ a posição $x = 0$ não é mais a de menor energia e torna-se instável. Há uma degenerescência dos estados de mais baixa energia ($x = \pm\sqrt{l^2 - a^2}$) e a simetria é quebrada, pois a transformação $x \to -x$ leva um estado de equilíbrio estável no outro. A Fig. 5.4 é um gráfico da energia E_0 do estado fundamental (estado de equilíbrio de mais baixa energia) em função da distância a. Há uma descontinuidade na segunda derivada de E_0 no ponto crítico $a = l$. A instabilidade em $x = 0$ quando a decresce e passa pelo valor crítico $a = l$ é análoga a uma transição de fase de segunda ordem em mecânica estatística, com o parâmetro a desempenhando o papel da temperatura. É surpreendente que este sistema mecânico tão simples seja capaz de revelar todas as características essenciais do fenômeno de quebra espontânea de simetria, uma das idéias centrais da moderna física das partículas elementares e ingrediente fundamental do modelo de unificação das interações eletrofracas.

5.2 Movimento Estacionário e Pequenas Oscilações

As considerações anteriores podem ser estendidas a certos sistemas pluridimensionais capazes de executar movimentos estacionários, que podem ser encarados como estados de equilíbrio dinâmico. Tais movimentos ocorrem tipicamente em sistemas que possuem coordenadas cíclicas. Adotaremos a seguinte definição: um movimento é dito estacionário se, em algum sistema de coordenadas, as coordenadas não-cíclicas, isto é, aquelas que aparecem na lagrangiana, são constantes.

■ **Exemplo 5.2.1.** Em coordenadas polares a lagrangiana de uma partícula movendo-se num plano sob uma força central é

$$L = \frac{m}{2}(r^2 + r^2\dot{\theta}^2) - V(r) \ . \tag{5.2.1}$$

A coordenada θ é cíclica, donde $mr^2\dot{\theta} = p_\theta =$ constante. A equação de movimento radial

$$m\ddot{r} - mr\dot{\theta}^2 + \frac{dV}{dr} = 0 \tag{5.2.2}$$

possui a solução estacionária $r = r_0$ (órbita circular) contanto que $(dV/dr)_{r=r_0} = p_\theta^2/(mr_0^3)$. Por exemplo, se $V(r) = -A/r$ com $A > 0$ temos $r_0 = p_\theta^2/(mA)$. Há, portanto, órbitas circulares no caso gravitacional ou no coulombiano atrativo. ■

■ **Exemplo 5.2.2.** A lagrangiana de um pêndulo esférico (isto é, não restrito a oscilar num plano fixo) é dada por

$$L = \frac{ml^2}{2}(\dot{\theta}^2 + \dot{\varphi}^2 \operatorname{sen}^2\theta) - mgl(1 - \cos\theta) \ . \tag{5.2.3}$$

A coordenada φ é cíclica, de modo que $ml^2\dot{\varphi}\operatorname{sen}^2\theta = p_\varphi = \text{constante}$. A equação para θ é

$$ml^2\ddot{\theta} - ml^2\dot{\varphi}^2 \operatorname{sen}\theta\cos\theta + mgl\operatorname{sen}\theta = 0 \ , \tag{5.2.4}$$

que admite a solução estacionária $\theta = \theta_0$ com $\cos\theta_0 = g/(l\dot{\varphi}^2)$ desde que $\dot{\varphi} > (g/l)^{1/2}$. ■

Nos dois exemplos anteriores a lagrangiana não depende explicitamente do tempo e a energia cinética é função homogênea do segundo grau das velocidades generalizadas, de sorte que $h = E$ e h é constante de movimento. Em ambos os casos a conservação da energia exprime-se na forma

$$\frac{\alpha(q)}{2}\dot{q}^2 + V_{\text{ef}}(q) = E \ , \tag{5.2.5}$$

onde q denota a coordenada não-cíclica. Com efeito, no Exemplo 5.2.1,

$$E = T + V = \frac{m}{2}\dot{r}^2 + \frac{m}{2}r^2\dot{\theta}^2 + V(r) = \frac{m}{2}\dot{r}^2 + \frac{p_\theta^2}{2mr^2} + V(r) \ , \tag{5.2.6}$$

ao passo que, no Exemplo 5.2.2, usando $ml^2\dot{\varphi}^2\operatorname{sen}^2\theta = p_\varphi$,

$$E = T + V = \frac{ml^2}{2}\dot{\theta}^2 + \frac{p_\varphi^2}{2ml^2\operatorname{sen}^2\theta} + mgl(1 - \cos\theta) \ . \tag{5.2.7}$$

As configurações de equilíbrio estacionário são as soluções de

$$\frac{dV_{\text{ef}}(q)}{dq} = 0 \ , \tag{5.2.8}$$

enquanto que a estabilidade depende do sinal da segunda derivada do *potencial efetivo* nas posições de equilíbrio, o sinal positivo implicando equilíbrio estável. É importante destacar que o potencial efetivo é identificado na equação que exprime a conservação da energia após a eliminação da velocidade associada à coordenada cíclica em favor do seu momento conjugado, que é constante de movimento.

■ **Exemplo 5.2.3.** Investigar a estabilidade da órbita circular de um corpo celeste em torno do Sol.

Solução. Segundo a Eq.(5.2.6),

$$V_{\text{ef}}(r) = \frac{p_\theta^2}{2mr^2} - \frac{A}{r} \tag{5.2.9}$$

com $A = GmM > 0$, sendo M a massa do Sol e m a massa do corpo celeste em órbita circular. O raio da órbita determina-se de

$$0 = \frac{dV_{\text{ef}}}{dr} = -\frac{p_\theta^2}{mr^3} + \frac{A}{r^2} \quad \Longrightarrow \quad r = r_0 = \frac{p_\theta^2}{mA} \ , \tag{5.2.10}$$

já obtido de outra forma no Exemplo 5.2.1. Por outro lado,

$$k^{(0)} = \left(\frac{d^2V_{\text{ef}}}{dr^2}\right)_{r=r_0} = \frac{3p_\theta^2}{mr_0^4} - \frac{2A}{r_0^3} = \frac{A}{r_0^3} > 0 \ , \tag{5.2.11}$$

e as órbitas circulares são estáveis frente a pequenas perturbações,[2] sendo

$$\omega = \left(\frac{k^{(0)}}{\alpha^{(0)}}\right)^{1/2} = \left(\frac{GM}{r_0^3}\right)^{1/2} \tag{5.2.12}$$

a freqüência das pequenas oscilações em torno da órbita circular. ■

■ **Exemplo 5.2.4.** Estudar a estabilidade do movimento estacionário de precessão ($\theta = \theta_0$) de um pêndulo esférico.

Solução. De (5.2.7) infere-se

$$V_{\text{ef}}(\theta) = mgl(1 - \cos\theta) + \frac{p_\varphi^2}{2ml^2\,\text{sen}^2\theta} \ . \tag{5.2.13}$$

O movimento estacionário $\theta = \theta_0$ requer

$$\left(\frac{dV_{\text{ef}}}{d\theta}\right)_{\theta=\theta_0} = mgl\,\text{sen}\,\theta_0 - \frac{p_\varphi^2\cos\theta_0}{ml^2\,\text{sen}^3\theta_0} = 0 \ . \tag{5.2.14}$$

Fazendo uso de $p_\varphi = ml^2\,\text{sen}^2\theta_0\,\dot\varphi$ resulta $\cos\theta_0 = g/(l\dot\varphi^2)$ contanto que $\dot\varphi > (g/l)^{1/2}$. Por outro lado,

$$k^{(0)} = \left(\frac{d^2V_{\text{ef}}}{d\theta^2}\right)_{\theta=\theta_0} = mgl\cos\theta_0 + \frac{p_\varphi^2}{ml^2\,\text{sen}^2\theta_0} + 3\frac{p_\varphi^2\cos^2\theta_0}{ml^2\,\text{sen}^4\theta_0} = \frac{mgl}{\cos\theta_0}(1 + 3\cos^2\theta_0) > 0 \ , \tag{5.2.15}$$

onde usamos (5.2.14). A precessão é estável e

$$\omega = \left(\frac{k^{(0)}}{\alpha^{(0)}}\right)^{1/2} = \left(\frac{k^{(0)}}{ml^2}\right)^{1/2} = \left(\frac{g}{l}\frac{1 + 3\cos^2\theta_0}{\cos\theta_0}\right)^{1/2} \tag{5.2.16}$$

é a freqüência das pequenas oscilações em torno do movimento estacionário. ■

[2] A presente análise não se aplica a perturbações perpendiculares ao plano da órbita.

5.3 Pequenas Oscilações: Caso Geral

Com as mesmas hipóteses do caso unidimensional, um sistema conservativo com n graus de liberdade descrito pelas coordenadas generalizadas $q_1, \ldots, q_n$ tem lagrangiana

$$L = \frac{1}{2} \sum_{k,l=1}^{n} M_{kl}\, \dot{q}_k\, \dot{q}_l - V(q_1, \ldots, q_n) \tag{5.3.1}$$

onde os $M_{kl} = M_{kl}(q_1, \ldots, q_n)$ são quantidades simétricas em seus índices, isto é, $M_{kl} = M_{lk}$ [vide Eq.(2.6.5c)]. A configuração $q^{(0)} \equiv (q_1^{(0)}, \ldots, q_n^{(0)})$ é uma configuração de equilíbrio se

$$\left(\frac{\partial V}{\partial q_k}\right)_{q^{(0)}} = 0 \quad , \quad k = 1, \ldots, n \ . \tag{5.3.2}$$

Definindo

$$q_k = q_k^{(0)} + \eta_k \quad , \quad k = 1, \ldots, n \ , \tag{5.3.3}$$

e expandindo $V(q)$ até segunda ordem em potências de $\eta_1, \ldots, \eta_n$ resulta

$$V(q_1, \ldots, q_n) = V(q^{(0)}) + \sum_k \left(\frac{\partial V}{\partial q_k}\right)_{q^{(0)}} \eta_k + \frac{1}{2} \sum_{kl} \left(\frac{\partial^2 V}{\partial q_k \partial q_l}\right)_{q^{(0)}} \eta_k\, \eta_l + \cdots \ . \tag{5.3.4}$$

O termo linear nos pequenos deslocamentos é nulo em virtude de (5.3.2) e a aproximação dominante para V toma a forma

$$V = V_0 + \frac{1}{2} \sum_{kl} V_{kl}\, \eta_k\, \eta_l \ , \tag{5.3.5}$$

onde $V_0 = V(q^{(0)})$ é constante e

$$V_{kl} = \left(\frac{\partial^2 V}{\partial q_k \partial q_l}\right)_{q^{(0)}} \ . \tag{5.3.6}$$

As quantidades V_{kl} são constantes e simétricas: $V_{kl} = V_{lk}$. O critério de estabilidade agora é um pouco mais complicado. Se V é mínimo em $q = q^{(0)}$ seu valor só pode aumentar em conseqüência de um pequeno deslocamento da configuração $q = q^{(0)}$. Pela Eq.(5.3.4), portanto, $q = q^{(0)}$ é ponto de equilíbrio estável se

$$\sum_{kl} V_{kl}\, \eta_k\, \eta_l > 0 \quad \forall\ (\eta_1, \ldots, \eta_n) \neq (0, \ldots, 0) \ . \tag{5.3.7}$$

Em resumo, o equilíbrio é estável se a *forma quadrática* (5.3.7) construída com as segundas derivadas do potencial é positiva definida, hipótese que faremos de agora em diante.

Assim como no problema unidimensional, a expansão da lagrangiana (5.3.1), correta até segunda ordem em η e $\dot{\eta}$, escreve-se

$$L = \frac{1}{2}\sum_{kl} T_{kl}\,\dot{\eta}_k\,\dot{\eta}_l - \frac{1}{2}\sum_{kl} V_{kl}\,\eta_k\,\eta_l \ , \tag{5.3.8}$$

onde

$$T_{kl} = M_{kl}(q^{(0)}) \tag{5.3.9}$$

e a constante aditiva V_0 foi descartada. As equações de movimento para os pequenos deslocamentos da posição de equilíbrio são

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\eta}_j}\right) - \frac{\partial L}{\partial \eta_j} = 0 \quad , \quad j = 1, \ldots, n \ . \tag{5.3.10}$$

Mas

$$\frac{\partial L}{\partial \dot{\eta}_j} = \frac{1}{2}\sum_{kl} T_{kl}\,(\delta_{kj}\,\dot{\eta}_l + \dot{\eta}_k\,\delta_{lj}) = \frac{1}{2}\Big\{\sum_{l} T_{jl}\,\dot{\eta}_l + \sum_{k} T_{kj}\,\dot{\eta}_k\Big\} = \frac{1}{2}\Big\{\sum_{l} T_{jl}\,\dot{\eta}_l + \sum_{k} T_{jk}\,\dot{\eta}_k\Big\} \ , \tag{5.3.11}$$

porque $T_{kj} = T_{jk}$. Finalmente, levando em conta que o índice de soma é "mudo", podemos trocar "l" por "k" no penúltimo somatório que aparece na Eq.(5.3.11), obtendo

$$\frac{\partial L}{\partial \dot{\eta}_j} = \sum_{k} T_{jk}\,\dot{\eta}_k \ . \tag{5.3.12}$$

Analogamente,

$$-\frac{\partial L}{\partial \eta_j} = \sum_{k} V_{jk}\,\eta_k \ , \tag{5.3.13}$$

de modo que as equações de Lagrange (5.3.10) assumem a forma

$$\sum_{k} T_{jk}\,\ddot{\eta}_k + \sum_{k} V_{jk}\,\eta_k = 0 \quad , \quad j = 1, \ldots, n \ , \tag{5.3.14}$$

que são equações de movimento de osciladores acoplados.

5.4 Modos Normais de Vibração

Como logo veremos, a solução geral das equações de movimento acopladas (5.3.14) pode ser expressa em termos de soluções particulares em que todos os deslocamentos $\eta_1, \ldots, \eta_n$ executam oscilações harmônicas com a mesma freqüência. Tais soluções particulares são conhecidas como *modos normais de vibração*. É conveniente substituir $\eta_k(t)$ pela função complexa $z_k(t)$ e definir

$$\eta_k(t) = \mathrm{Re}\, z_k(t) \quad , \quad k = 1, \ldots, n \ . \tag{5.4.1}$$

Como os coeficientes do sistema de equações (5.3.14) são reais, as partes real e imaginária de uma solução complexa são também soluções, o que justifica a Eq.(5.4.1).

Freqüências Características

Busquemos soluções complexas de (5.3.14) (com $z_k(t)$ no lugar de $\eta_k(t)$) na forma

$$z_k(t) = z_k^{(0)} e^{i\omega t} \quad , \quad k = 1, \ldots, n \ , \tag{5.4.2}$$

em que todos os deslocamentos vibram com a mesma freqüência ω. Introduzindo a expressão acima no lugar de $\eta_k(t)$ em (5.3.14) resulta

$$\sum_k (V_{jk} - \omega^2 T_{jk})\, z_k^{(0)} = 0 \ . \tag{5.4.3}$$

Definindo as matrizes

$$\mathbf{V} = (V_{jk}) \quad , \quad \mathbf{T} = (T_{jk}) \quad , \quad \mathbf{z}^{(0)} = \begin{pmatrix} z_1^{(0)} \\ \vdots \\ z_n^{(0)} \end{pmatrix} \ , \tag{5.4.4}$$

as equações (5.4.3) equivalem à equação matricial

$$(\mathbf{V} - \omega^2 \mathbf{T})\, \mathbf{z}^{(0)} = 0 \ . \tag{5.4.5}$$

Esta última equação só possui solução não-trivial $\mathbf{z}^{(0)} \neq 0$ se

$$\det(\mathbf{V} - \omega^2 \mathbf{T}) = \det \begin{pmatrix} V_{11} - \omega^2 T_{11} & V_{12} - \omega^2 T_{12} & \cdots \\ V_{21} - \omega^2 T_{21} & V_{22} - \omega^2 T_{22} & \cdots \\ \vdots & \vdots & \vdots \end{pmatrix} = 0 \ , \tag{5.4.6}$$

que é uma equação algébrica de grau n para ω^2, possuindo n raízes positivas

$$\omega = \omega_s \quad , \quad s = 1, \ldots, n \ , \tag{5.4.7}$$

onde as quantidades positivas $\omega_1, \ldots, \omega_n$ são chamadas de *freqüências características* do sistema.

Modos Normais de Vibração

É possível encontrar n vetores linearmente independentes $\boldsymbol{\varrho}^{(1)}, \ldots, \boldsymbol{\varrho}^{(n)}$, chamados de *vetores característicos*, que satisfazem (5.4.5), isto é,

$$(\mathbf{V} - \omega_{(s)}^2 \mathbf{T})\, \boldsymbol{\varrho}^{(s)} = 0 \quad , \quad s = 1, \ldots, n \ . \tag{5.4.8}$$

Como $\mathbf{V}$ e $\mathbf{T}$ são matrizes reais, é sempre possível escolher cada $\boldsymbol{\varrho}^{(s)}$ de modo que todas

as suas componentes sejam reais. Supondo que tal escolha tenha sido feita, a solução complexa associada à freqüência ω_s escreve-se

$$\mathbf{z}^{(s)}(t) = \boldsymbol{\varrho}^{(s)}\, e^{i\omega_s t} \ . \tag{5.4.9}$$

Cada solução da forma (5.4.9) é um modo normal de vibração do sistema.

Solução Geral das Equações de Movimento

As Eqs.(5.3.14) são lineares, de modo que qualquer combinação linear de soluções também é solução. Assim, a solução mais geral possível para $\mathbf{z}(t)$ é expressa pela superposição das soluções associadas às freqüências características, isto é,

$$\mathbf{z}(t) = \sum_{s=1}^{n} A^{(s)}\, \boldsymbol{\varrho}^{(s)}\, e^{i\omega_s t} \tag{5.4.10}$$

com coeficientes complexos arbitrários $A^{(s)}$. Escrevendo $A^{(s)}$ na forma polar

$$A^{(s)} = C^{(s)}\, e^{i\phi_s} \tag{5.4.11}$$

com $C^{(s)}$ e ϕ_s reais, ficamos com

$$\mathbf{z}(t) = \sum_{s=1}^{n} C^{(s)}\, \boldsymbol{\varrho}^{(s)}\, e^{i\,(\omega_s t + \phi_s)} \ . \tag{5.4.12}$$

Uma vez que os $\boldsymbol{\varrho}^{(s)}$ são vetores reais por hipótese, a solução física é

$$\boldsymbol{\eta}(t) = \mathrm{Re}\,\mathbf{z}(t) = \sum_{s=1}^{n} C^{(s)}\, \boldsymbol{\varrho}^{(s)}\, \cos\left(\omega_s t + \phi_s\right) \tag{5.4.13a}$$

ou, em componentes,

$$\eta_k(t) = \sum_{s=1}^{n} C^{(s)}\, \varrho_k^{(s)}\, \cos\left(\omega_s t + \phi_s\right) \quad , \quad k = 1, \ldots, n \ . \tag{5.4.13b}$$

A solução geral das equações de movimento (5.3.14), expressa pela última equação acima, contém as $2n$ constantes arbitrárias $C^{(1)}, \ldots, C^{(n)}, \phi_1, \ldots, \phi_n$ que podem ser determinadas pelas condições iniciais.

■ **Exemplo 5.4.1.** Dois pêndulos iguais executam pequenas oscilações acoplados por uma mola de massa desprezível. Determinar os modos normais de vibração e a solução das equações de movimento se, em $t = 0$, um dos pêndulos está ligeiramente deslocado de sua posição de equilíbrio e o movimento tem início com ambos os pêndulos em repouso.

Solução. Suponhamos que a configuração de equilíbrio seja $\theta_1 = \theta_2 = 0$, na qual a mola possui o seu comprimento natural d_0. Usaremos como coordenadas os pequenos deslocamentos horizontais η_1 e η_2 indicados na Fig. 5.5, que relacionam-se com os pequenos ângulos θ_1 e θ_2 por

$$\eta_1 = l \,\mathrm{sen}\,\theta_1 \approx l\,\theta_1 \quad , \quad \eta_2 = l \,\mathrm{sen}\,\theta_2 \approx l\,\theta_2 \ . \tag{5.4.14}$$

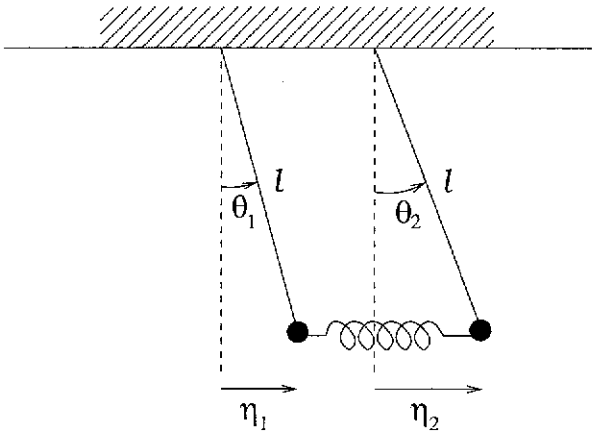


Figura 5.5: Dois pêndulos iguais acoplados por uma mola.

A energia cinética é

$$T = \frac{ml^2}{2}\dot{\theta}_1^2 + \frac{ml^2}{2}\dot{\theta}_2^2 \approx \frac{m}{2}\dot{\eta}_1^2 + \frac{m}{2}\dot{\eta}_2^2 \ , \tag{5.4.15}$$

ao passo que a energia potencial escreve-se

$$V = \frac{k}{2}\left[(d_0 + \eta_2 - \eta_1) - d_0\right]^2 + mg[l(1-\cos\theta_1) + l(1-\cos\theta_2)] \ . \tag{5.4.16}$$

Usando $\cos\theta \approx 1 - \theta^2/2$ e (5.4.14), resulta

$$V = \frac{k}{2}(\eta_2 - \eta_1)^2 + \frac{mg}{2l}({\eta_1}^2 + {\eta_2}^2) \ , \tag{5.4.17}$$

de modo que

$$L = \frac{m}{2}(\dot{\eta}_1^2 + \dot{\eta}_2^2) - \frac{1}{2}\left\{\frac{mg}{l}({\eta_1}^2 + {\eta_2}^2) + k(\eta_2 - \eta_1)^2\right\} \ . \tag{5.4.18}$$

Comparando com a forma padrão (5.3.8), inferimos

$$\mathbf{T} = \begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix} \quad , \quad \mathbf{V} = \begin{pmatrix} \frac{mg}{l} + k & -k \\ -k & \frac{mg}{l} + k \end{pmatrix} \ . \tag{5.4.19}$$

Advertência. Na forma padrão (5.3.8) cada termo cruzado aparece duas vezes. Por exemplo, na segunda soma dupla sobre k e l há um termo com $k = 1, l = 2$ e outro com $k = 2, l = 1$ que são iguais entre si porque $V_{12} = V_{21}$. Assim, o coeficiente de $\eta_1\eta_2$ é $V_{12} + V_{21} = 2V_{12}$. Por outro lado, o coeficiente de $\eta_1\eta_2$ na expressão entre chaves em (5.4.18) é $-2k$, daí a identificação $V_{12} = -k$. Este tipo de cuidado deve sempre ser tomado ao se comparar a forma padrão (5.3.8) com uma lagrangiana específica tal como (5.4.18), caso contrário erros por um fator dois serão cometidos que comprometerão todos os resultados subseqüentes.

As freqüências dos modos normais são as soluções de

$$\det(\mathbf{V} - \omega^2\mathbf{T}) = \det\begin{pmatrix} \frac{mg}{l} + k - m\omega^2 & -k \\ -k & \frac{mg}{l} + k - m\omega^2 \end{pmatrix} = 0 \ , \tag{5.4.20}$$

ou seja,

$$\left(\frac{mg}{l}+k-m\omega^2\right)^2-k^2=0 \quad\Longrightarrow\quad \frac{mg}{l}+k-m\omega^2=\pm k \ . \tag{5.4.21}$$

As freqüências características são

$$\omega_1=\left(\frac{g}{l}\right)^{1/2} \quad , \quad \omega_2=\left(\frac{g}{l}+\frac{2k}{m}\right)^{1/2} \ , \tag{5.4.22}$$

e os autovetores $\varrho^{(s)}$ correspondentes são tais que

$$(\mathbf{V}-\omega_1^2\mathbf{T})\,\varrho^{(1)}=0 \Longrightarrow \begin{pmatrix} k & -k \\ -k & k \end{pmatrix}\begin{pmatrix} \varrho_1^{(1)} \\ \varrho_2^{(1)} \end{pmatrix}=0 \Longrightarrow \varrho_1^{(1)}=\varrho_2^{(1)} \Longrightarrow \varrho^{(1)}=\begin{pmatrix} \alpha \\ \alpha \end{pmatrix}, \tag{5.4.23a}$$

$$(\mathbf{V}-\omega_2^2\mathbf{T})\,\varrho^{(2)}=0 \Longrightarrow \begin{pmatrix} -k & -k \\ -k & -k \end{pmatrix}\begin{pmatrix} \varrho_1^{(2)} \\ \varrho_2^{(2)} \end{pmatrix}=0 \Longrightarrow \varrho_1^{(2)}=-\varrho_2^{(2)} \Longrightarrow \varrho^{(2)}=\begin{pmatrix} \beta \\ -\beta \end{pmatrix}, \tag{5.4.23b}$$

onde α e β são números reais arbitrários. O modo de freqüência ω_1 corresponde a um movimento oscilatório dos pêndulos com $\eta_1=\eta_2$. Neste caso, tudo se passa como se a mola não existisse, pois ela permanece todo o tempo com o seu comprimento natural, o que explica o valor de ω_1 dado em (5.4.22): ambos os pêndulos oscilam concertadamente com a freqüência de um pêndulo simples isolado. O segundo modo normal consiste num movimento oscilatório de ambos os pêndulos com freqüência ω_2, porém agora $\eta_1=-\eta_2$, conforme a Fig. 5.6.

As duas soluções reais e linearmente independentes podem ser escolhidas como

$$\varrho^{(1)}=\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad , \quad \varrho^{(2)}=\begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{5.4.24}$$

e a solução geral (5.4.13) toma a forma

$$\eta_1(t)=\sum_{s=1}^{2} C^{(s)}\,\varrho_1^{(s)}\cos\left(\omega_s t+\phi_s\right)=C^{(1)}\cos\left(\omega_1 t+\phi_1\right)+C^{(2)}\cos\left(\omega_2 t+\phi_2\right) \ , \tag{5.4.25a}$$

$$\eta_2(t)=\sum_{s=1}^{2} C^{(s)}\,\varrho_2^{(s)}\cos\left(\omega_s t+\phi_s\right)=C^{(1)}\cos\left(\omega_1 t+\phi_1\right)-C^{(2)}\cos\left(\omega_2 t+\phi_2\right) \ . \tag{5.4.25b}$$

Das condições iniciais $\eta_1(0)=0, \eta_2(0)=a, \dot{\eta}_1(0)=\dot{\eta}_2(0)=0$ inferimos

$$C^{(1)}\cos\phi_1+C^{(2)}\cos\phi_2=0 \ , \tag{5.4.26a}$$

$$C^{(1)}\cos\phi_1-C^{(2)}\cos\phi_2=a \ , \tag{5.4.26b}$$

$$-\omega_1 C^{(1)}\,\text{sen}\,\phi_1-\omega_2 C^{(2)}\,\text{sen}\,\phi_2=0 \ , \tag{5.4.26c}$$

$$-\omega_1 C^{(1)}\,\text{sen}\,\phi_1+\omega_2 C^{(2)}\,\text{sen}\,\phi_2=0 \ . \tag{5.4.26d}$$

Somando as duas primeiras e as duas últimas das Eqs.(5.4.26), deduzimos

$$C^{(1)}\cos\phi_1=-C^{(2)}\cos\phi_2=a/2 \ , \tag{5.4.27a}$$

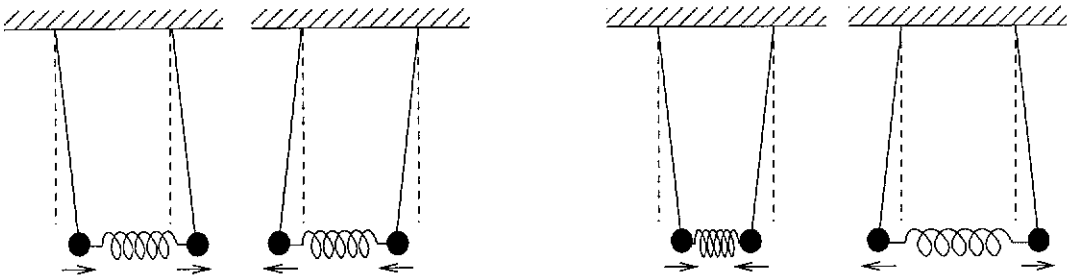


Figura 5.6: Modos normais de vibração de dois pêndulos iguais acoplados por uma mola.

$$C^{(1)} \text{ sen } \phi_1 = C^{(2)} \text{ sen } \phi_2 = 0 \quad . \tag{5.4.27b}$$

Não é admissível ter-se $C^{(1)} = 0$ ou $C^{(2)} = 0$ porque isto geraria uma contradição com (5.4.27a). Logo, sen $\phi_1 =$ sen $\phi_2 = 0$, cuja solução mais simples é $\phi_1 = \phi_2 = 0$, em decorrência da qual $C^{(1)} = -C^{(2)} = a/2$. Finalmente, a solução das equações de movimento escreve-se

$$\eta_1(t) = \frac{a}{2}(\cos \omega_1 t - \cos \omega_2 t) \ , \tag{5.4.28a}$$

$$\eta_2(t) = \frac{a}{2}(\cos \omega_1 t + \cos \omega_2 t) \ , \tag{5.4.28b}$$

com ω_1 e ω_2 dadas por (5.4.22). ■

5.5 Coordenadas Normais

A existência de n modos normais de vibração com freqüências $\omega_1, \ldots, \omega_n$ associados à lagrangiana (5.3.8) sugere que ela seja equivalente à lagrangiana de um conjunto de osciladores harmônicos independentes, cada um deles com freqüência igual a uma das freqüências características. Isto, no entanto, não está explícito na lagrangiana (5.3.8), pois as coordenadas $\eta_1, \ldots, \eta_n$ estão misturadas umas com as outras, resultando nas equações de movimento acopladas (5.3.14). Deve ser possível introduzir novas coordenadas em lugar de $\eta_1, \ldots, \eta_n$ de modo a eliminar todos os termos cruzados na lagrangiana (5.3.8), reduzindo-a a uma soma de lagrangianas de osciladores harmônicos independentes. Tais coordenadas sempre existem, e passaremos a expor como elas podem ser obtidas em geral. A fim de permitir que a exposição flua suavemente, sem constantes interrupções, as demonstrações das hipóteses ou afirmações matemáticas feitas na Seção 5.3, e outras que necessariamente surgirão ao longo do texto abaixo, serão realizadas na próxima seção.

Introduzindo a notação matricial

$$\boldsymbol{\eta} = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} \quad , \quad \boldsymbol{\eta}^T = (\eta_1 \ \ldots \ \eta_n) \ , \tag{5.5.1}$$

a lagrangiana (5.3.8) escreve-se

$$L = \frac{1}{2}\dot{\eta}^T \mathbf{T}\dot{\eta} - \frac{1}{2}\eta^T \mathbf{V}\eta \ . \tag{5.5.2}$$

Defina novas coordenadas $\zeta_1, \ldots, \zeta_n$ relacionadas com $\eta_1, \ldots, \eta_n$ pela transformação linear

$$\eta = \mathbf{A}\zeta \ , \tag{5.5.3}$$

onde $\mathbf{A}$ é uma matriz a ser determinada. Substituindo (5.5.3) em (5.5.2) resulta

$$L = \frac{1}{2}\dot{\zeta}^T \mathbf{A}^T \mathbf{T}\mathbf{A}\dot{\zeta} - \frac{1}{2}\zeta^T \mathbf{A}^T \mathbf{V}\mathbf{A}\zeta \ . \tag{5.5.4}$$

Para que a lagrangiana não contenha acoplamentos entre as coordenadas $\zeta_1, \ldots, \zeta_n$ é necessário que as matrizes $\mathbf{A}^T\mathbf{T}\mathbf{A}$ e $\mathbf{A}^T\mathbf{V}\mathbf{A}$ sejam ambas diagonais. O problema, portanto, consiste na busca de uma matriz $\mathbf{A}$ capaz de diagonalizar simultaneamente as matrizes $\mathbf{T}$ e $\mathbf{V}$.

A introdução de um produto interno no espaço vetorial formado pelas matrizes-coluna de n componentes reais facilita a construção da matriz $\mathbf{A}$ com os requisitos desejados. Defina, portanto, o produto interno (η, ξ) por

$$(\eta, \xi) = \eta^T \mathbf{T}\xi \equiv \sum_{kl} T_{kl}\,\eta_k\,\xi_l \ . \tag{5.5.5}$$

Se $\varrho^{(1)}, \ldots, \varrho^{(n)}$ são soluções linearmente independentes de (5.4.8), imponhamos a esses vetores a exigência adicional de que eles sejam escolhidos mutuamente ortogonais e normalizados segundo o produto interno recém-definido:

$$(\varrho^{(r)}, \varrho^{(s)}) = \delta_{rs} \quad , \quad r, s = 1, \ldots, n \ . \tag{5.5.6}$$

Defina, agora, a *matriz modal* $\mathbf{A} = (A_{kr})$ por

$$A_{kr} = \varrho_k^{(r)} \ , \tag{5.5.7}$$

donde[3]

$$\mathbf{A} = \begin{pmatrix} \varrho^{(1)} & \varrho^{(2)} & \cdots & \varrho^{(n)} \\ \downarrow & \downarrow & \cdots & \downarrow \end{pmatrix} \ , \tag{5.5.8}$$

isto é, a p-ésima coluna de $\mathbf{A}$ é o vetor $\varrho^{(p)}$. As equações (5.5.6) equivalem a

$$\sum_{kl} \varrho_k^{(r)}\, T_{kl}\, \varrho_l^{(s)} = \delta_{rs} \ , \tag{5.5.9}$$

que ainda podem ser reescritas na forma

[3] Aqui adotamos a notação e a terminologia de Fetter & Walecka (1980).

$$\sum_k (\mathbf{A}^T)_{rk} \left(\sum_l T_{kl}\, A_{ls}\right) = \sum_k (\mathbf{A}^T)_{rk}\, (\mathbf{TA})_{ks} = \delta_{rs} \ , \tag{5.5.10}$$

Usando novamente a definição de produto de matrizes somos conduzidos a

$$\mathbf{A}^T \mathbf{TA} = \boldsymbol{I} \ . \tag{5.5.11}$$

A matriz modal (5.5.8) reduz a matriz $\mathbf{T}$ à forma diagonal mais simples possível – a matriz identidade. Resta verificar se a matriz $\mathbf{V}$ é também reduzida à forma diagonal.

Em componentes, as equações (5.4.8) escrevem-se

$$\sum_l V_{kl}\, \varrho_l^{(s)} = \omega_s^2 \sum_l T_{kl}\, \varrho_l^{(s)} \tag{5.5.12}$$

ou, tirando partido de (5.5.7),

$$(\mathbf{VA})_{ks} = \omega_s^2\, (\mathbf{TA})_{ks} \ . \tag{5.5.13}$$

Defina a matriz diagonal $\mathbf{W}^D = (W_{rs})$ por

$$W_{rs} = \omega_s^2\, \delta_{rs} \ , \tag{5.5.14}$$

isto é,

$$\mathbf{W}^D = \begin{pmatrix} \omega_1^2 & 0 & \dots & 0 \\ 0 & \omega_2^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \omega_n^2 \end{pmatrix} \ . \tag{5.5.15}$$

A Eq.(5.5.13) pode ser recscrita na forma equivalente

$$(\mathbf{VA})_{ks} = \sum_r (\mathbf{TA})_{kr} \omega_s^2\, \delta_{rs} = \sum_r (\mathbf{TA})_{kr}\, (\mathbf{W}^D)_{rs} \ , \tag{5.5.16}$$

ou seja,

$$\mathbf{VA} = \mathbf{TA}\mathbf{W}^D \ . \tag{5.5.17}$$

Multiplicando esta equação pela esquerda por $\mathbf{A}^T$ e utilizando (5.5.11) obtém-se, finalmente,

$$\mathbf{A}^T \mathbf{VA} = \mathbf{W}^D \ , \tag{5.5.18}$$

completando a demonstração de que $\mathbf{A}$ também reduz $\mathbf{V}$ à forma diagonal.

Em termos das *coordenadas normais* $\zeta_1, \dots \zeta_n$ definidas por (5.5.3), a lagrangiana (5.5.2) toma a forma simplificada

$$L = \frac{1}{2}\dot{\zeta}^T\dot{\zeta} - \frac{1}{2}\zeta^T\mathbf{W}^D\zeta = \sum_{k=1}^{n}\frac{1}{2}(\dot{\zeta}_k^2 - \omega_k^2\zeta_k^2) \ , \tag{5.5.19}$$

que descreve um conjunto de n osciladores harmônicos independentes com freqüências angulares $\omega_1, \ldots, \omega_n$, respectivamente. A equação de movimento para cada coordenada normal é simplesmente

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\zeta}_j}\right) - \frac{\partial L}{\partial \zeta_j} = 0 \quad \Longrightarrow \quad \ddot{\zeta}_j + \omega_j^2\zeta_j = 0 \quad , \quad j = 1, \ldots, n \ , \tag{5.5.20}$$

cuja solução geral é

$$\zeta_j(t) = C^{(j)}\cos(\omega_j t + \phi_j) \quad , \quad j = 1, \ldots, n \ . \tag{5.5.21}$$

Cada modo normal de vibração consiste num movimento do sistema em que apenas uma das coordenadas normais oscila harmonicamente, todas as demais permanecendo identicamente nulas. Finalmente, a expressão das coordenadas normais em função das coordenadas originais $\eta_1, \ldots, \eta_n$ pode ser imediatamente encontrada por inversão da Eq.(5.5.3): multiplique-a pela esquerda por $\mathbf{A}^T\mathbf{T}$ e use (5.5.11) para obter

$$\zeta = \mathbf{A}^T\mathbf{T}\eta \ . \tag{5.5.22}$$

■ **Exemplo 5.5.1.** Determinar as coordenadas normais para os pêndulos acoplados do Exemplo 5.4.1.

Solução. As duas soluções linearmente independentes das Eqs.(5.4.8) no caso dos pêndulos acoplados são

$$\varrho^{(1)} = \alpha\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad , \quad \varrho^{(2)} = \beta\begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad , \quad \alpha, \beta \in \boldsymbol{R} \ , \tag{5.5.23}$$

obtidas no Exemplo 5.4.1. A fim de construir a matriz modal é necessário que $\varrho^{(1)}$ e $\varrho^{(2)}$ constituam um conjunto ortonormal segundo o produto escalar definido pela matriz $\mathbf{T} = m\boldsymbol{I}$, isto é,

$$\varrho^{(1)^T}\mathbf{T}\varrho^{(1)} = 1 \implies m\alpha^2\begin{pmatrix} 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \implies 2m\alpha^2 = 1 \ , \tag{5.5.24a}$$

$$\varrho^{(2)^T}\mathbf{T}\varrho^{(2)} = 1 \implies m\beta^2\begin{pmatrix} 1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1 \implies 2m\beta^2 = 1 \ , \tag{5.5.24b}$$

$$\varrho^{(1)^T}\mathbf{T}\varrho^{(2)} = 0 \implies m\alpha\beta\begin{pmatrix} 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0 \implies m\alpha\beta(1-1) = 0 \ . \tag{5.5.24c}$$

Portanto, a condição de ortogonalidade é identicamente satisfeita, ao passo que a escolha $\alpha = \beta = (2m)^{-1/2}$ garante a normalização dos vetores, dando lugar a um conjunto ortonormal de soluções de (5.4.8) na forma

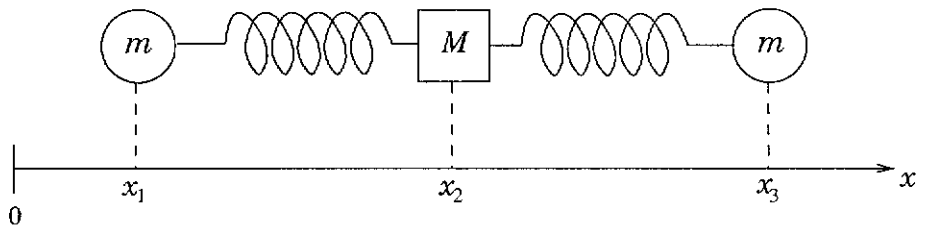


Figura 5.7: Modelo simplificado para o estudo das vibrações de uma molécula triatômica linear e simétrica.

$$\varrho^{(1)} = \frac{1}{\sqrt{2m}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad , \quad \varrho^{(2)} = \frac{1}{\sqrt{2m}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \ , \tag{5.5.25}$$

resultando na matriz modal

$$\mathbf{A} = \frac{1}{\sqrt{2m}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \ . \tag{5.5.26}$$

As coordenadas normais são

$$\zeta = \mathbf{A}^T \mathbf{T} \eta = \frac{m}{\sqrt{2m}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} \ , \tag{5.5.27}$$

isto é,

$$\zeta_1 = \sqrt{\frac{m}{2}} \, (\eta_1 + \eta_2) \quad , \quad \zeta_2 = \sqrt{\frac{m}{2}} \, (\eta_1 - \eta_2) \ . \tag{5.5.28}$$

O modo normal de freqüência ω_1 tem $\eta_1 = \eta_2$, de maneira que somente ζ_1 oscila, ζ_2 sendo identicamente nula; o oposto ocorre com o modo normal de freqüência ω_2. ■

Como última ilustração, consideremos um sistema de interesse em espectroscopia molecular.

■ **Exemplo 5.5.2** Determinar os modos normais de vibração de uma molécula triatômica linear e simétrica (por exemplo, CO_2) representada pelo modelo da Fig. 5.7.

Solução. A energia potencial é dada por

$$V = \frac{k}{2} \, (x_2 - x_1 - \ell)^2 + \frac{k}{2} \, (x_3 - x_2 - \ell)^2 \ . \tag{5.5.29}$$

onde ℓ denota o comprimento natural de cada mola. A configuração de equilíbrio é tal que

$$0 = \frac{\partial V}{\partial x_1} = -k \, (x_2 - x_1 - \ell) \quad , \quad 0 = \frac{\partial V}{\partial x_2} = k \, (x_2 - x_1 - \ell) - k \, (x_3 - x_2 - \ell) \ , \tag{5.5.30}$$

donde

$$x_2 - x_1 = \ell \quad , \quad x_3 - x_2 = \ell \ , \tag{5.5.31}$$

como seria de se esperar. Para tornar nosso formalismo aplicável ao presente problema, é necessário eliminar da energia potencial termos lineares nas coordenadas generalizadas. Isto pode ser conseguido definindo

$$\eta_1 = x_1 + \ell \quad , \quad \eta_2 = x_2 \quad , \quad \eta_3 = x_3 - \ell \ , \tag{5.5.32}$$

de modo que

$$L = \frac{m}{2}(\dot{x}_1^2 + \dot{x}_3^2) + \frac{M}{2}\dot{x}_2^2 - V = \frac{m}{2}(\dot{\eta}_1^2 + \dot{\eta}_3^2) + \frac{M}{2}\dot{\eta}_2^2 - \frac{k}{2}\left[(\eta_2 - \eta_1)^2 + (\eta_3 - \eta_2)^2\right] \ . \tag{5.5.33}$$

Por inspeção identificam-se

$$\mathbf{T} = \begin{pmatrix} m & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & m \end{pmatrix} \quad , \quad \mathbf{V} = \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix} \ . \tag{5.5.34}$$

A equação característica escreve-se (pondo $\lambda = \omega^2$)

$$\det(\mathbf{V} - \lambda\mathbf{T}) = \det \begin{pmatrix} k - m\lambda & -k & 0 \\ -k & 2k - M\lambda & -k \\ 0 & -k & k - m\lambda \end{pmatrix} = 0 \ , \tag{5.5.35}$$

donde

$$(k - m\lambda)\left[(2k - M\lambda)(k - m\lambda) - k^2\right] - k^2\,(k - m\lambda) \ = \ 0 \ . \tag{5.5.36}$$

Uma raiz imediata é $\lambda = k/m$ e as demais são as soluções de

$$m\,M\,\lambda^2 - k\,(2m + M)\,\lambda \ = \ 0 \ , \tag{5.5.37}$$

de sorte que as freqüências dos modos normais são

$$\omega_1 = 0 \quad , \quad \omega_2 = \sqrt{\frac{k}{m}} \quad , \quad \omega_3 = \sqrt{\frac{k}{m}\left(1 + \frac{2m}{M}\right)} \ . \tag{5.5.38}$$

A primeira freqüência característica à primeira vista pode parecer estranha, pois não corresponde a movimento oscilatório, a coordenada normal associada obedecendo à equação $\ddot{\zeta}_1 = 0$. Esta freqüência nula deve-se à possibilidade de a molécula transladar-se rigidamente sem mudança na energia potencial. Como a força restauradora contra tal deslocamento é nula, a freqüência correspondente é zero. Matematicamente, isto corresponde ao fato de a forma quadrática associada à energia potencial não ser positiva definida. Com efeito, $V = 0$ para $\eta_1 = \eta_2 = \eta_3 \neq 0$, indicando que a energia potencial[4] não é mínima na posição de equilíbrio, o qual não é estável mas indiferente: um deslocamento rígido da molécula não está sujeito a nenhuma força tendente a restaurar o equilíbrio nem a afastar a molécula de sua configuração original. A amplitude $\varrho^{(1)}$ associada à freqüência $\omega_1 = 0$ satisfaz

[4] Uma forma quadrática não-negativa que se anula para valores diferentes de zero das variáveis é dita *positiva semidefinida*.

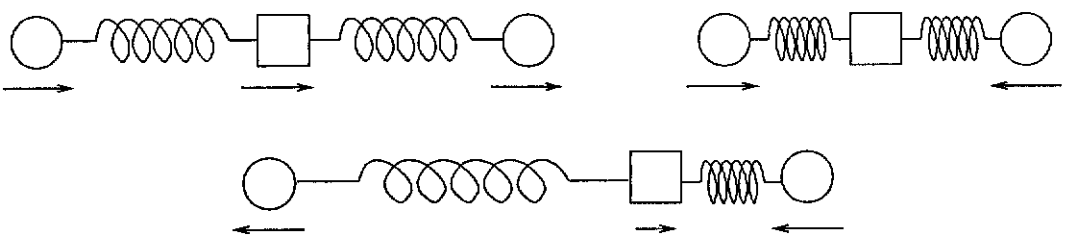


Figura 5.8: Modos normais de vibração da molécula triatômica linear e simétrica.

$$(\mathbf{V}-\lambda\mathbf{T})\,\varrho^{(1)}=0 \implies \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix}\begin{pmatrix} \varrho_1^{(1)} \\ \varrho_2^{(1)} \\ \varrho_3^{(1)} \end{pmatrix}=0 \implies \varrho^{(1)}=A\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{5.5.39}$$

Procedendo de forma análoga, encontramos

$$\varrho^{(2)}=B\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad , \quad \varrho^{(3)}=C\begin{pmatrix} 1 \\ -2m/M \\ 1 \end{pmatrix}. \tag{5.5.40}$$

Os modos normais de vibração estão representados esquematicamente na Fig. 5.8. Somente os modos 2 e 3 correspondem a vibrações da molécula, o modo número 1 consistindo numa translação rígida uniforme. Os vetores $\varrho^{(k)}$ estão associados a freqüências distintas, logo são automaticamente ortogonais no produto interno (5.5.5), como será demonstrado genericamente na próxima seção. A condição de normalização é

$$\varrho^{(s)^T}\mathbf{T}\varrho^{(s)}=\sum_{k\ell}T_{k\ell}\varrho_k^{(s)}\varrho_\ell^{(s)}=1 \Rightarrow m(\varrho_1^{(s)^2}+\varrho_3^{(s)^2})+M\varrho_2^{(s)^2}=1, \quad s=1,2,3 \tag{5.5.41}$$

de modo que os vetores normalizados podem ser escolhidos como

$$\varrho^{(1)}=\frac{1}{\sqrt{2m+M}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} , \; \varrho^{(2)}=\frac{1}{\sqrt{2m}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} , \; \varrho^{(3)}=\frac{1}{\sqrt{2m(1+2m/M)}}\begin{pmatrix} 1 \\ -2m/M \\ 1 \end{pmatrix} \tag{5.5.42}$$

Com a ajuda da matriz modal

$$\mathbf{A}=\begin{pmatrix} \varrho^{(1)} & \varrho^{(2)} & \varrho^{(3)} \\ \downarrow & \downarrow & \downarrow \end{pmatrix}, \tag{5.5.43}$$

as coordenadas normais escrevem-se $\boldsymbol{\zeta}=\mathbf{A}^T\mathbf{T}\boldsymbol{\eta}$, donde, por exemplo,

$$\zeta_1=\frac{m\eta_1+M\eta_2+m\eta_3}{\sqrt{2m+M}}=\frac{mx_1+Mx_2+mx_3}{\sqrt{2m+M}}=(2m+M)^{1/2}\,x_{\mathrm{CM}}\,. \tag{5.5.44}$$

Não havendo forças externas, o centro de massa move-se com velocidade constante ($\ddot{x}_{CM} = 0$), justificando a freqüência $\omega_1 = 0$ associada a ζ_1. ■

Outros exemplos mais complexos de vibrações moleculares podem ser encontrados em Landau & Lifchitz (1966). Para uma discussão detalhada e geral do problema das oscilações amortecidas e forçadas, que não trataremos aqui, uma excelente referência é Gantmacher (1970).

5.6 Suplemento Matemático

Nesta seção apresentamos as demonstrações que, para evitar a quebra da continuidade da exposição, foram omitidas nas duas seções anteriores. Aproveitamos para chamar a atenção também para a validade em geral de certas propriedades encontradas nos exemplos escolhidos para ilustrar a teoria. Sempre que necessário recorreremos ao Apêndice C, com o qual supomos que o leitor esteja familiarizado.

Energia Cinética e Produto Interno

A energia cinética de qualquer sistema físico é sempre positiva, e só é zero se todas a velocidades forem nulas, propriedade que, matematicamente, exprime-se por

$$\sum_{kl} T_{kl}\,\xi_k\,\xi_l \geq 0 \quad \forall \xi_1, \ldots, \xi_n \in \boldsymbol{R} \text{ e } \sum_{kl} T_{kl}\,\xi_k\,\xi_l = 0 \iff \xi_1 = \cdots = \xi_n = 0 \ . \tag{5.6.1}$$

Esta propriedade de positividade e a simetria da matriz $\mathbf{T}$, isto é, $T_{kl} = T_{lk}$, permitem demonstrar que (5.5.5) tem todas as propriedades de um produto interno, a saber:

$$(i)\ (\eta, \alpha\boldsymbol{\xi} + \beta\boldsymbol{\xi}') = \alpha(\eta, \boldsymbol{\xi}) + \beta(\eta, \boldsymbol{\xi}') \quad , \quad \alpha, \beta \in \boldsymbol{R} \qquad \text{(linearidade)} \ ; \tag{5.6.2a}$$

$$(ii)\ (\eta, \boldsymbol{\xi}) = (\boldsymbol{\xi}, \eta) \qquad \text{(simetria)} \ ; \tag{5.6.2b}$$

$$(iii)\ (\eta, \eta) \geq 0 \text{ e } (\eta, \eta) = 0 \iff \eta = 0 \qquad \text{(positividade)} \ . \tag{5.6.2c}$$

■ **Exercício 5.6.1.** Prove que $(\eta, \boldsymbol{\xi})$ definido em $\boldsymbol{R}^n$ por (5.5.5) goza das propriedades $(i) - (iii)$ que caracterizam um produto interno. ■

Positividade das Freqüências Características

Tacitamente admitimos que todas as soluções para ω^2 da equação característica (5.4.6) são positivas, acarretando que as freqüências $\omega_1, \ldots, \omega_n$ são reais e positivas. Isto pode ser facilmente demonstrado explorando a positividade da matriz $\mathbf{V}$. Seja $\lambda = \omega^2$ e suponha que

$$\mathbf{V}\boldsymbol{\xi} = \lambda \mathbf{T}\boldsymbol{\xi} \tag{5.6.3}$$

com $\boldsymbol{\xi} \neq 0$. Multiplicando esta última equação por $\boldsymbol{\xi}^T$ e resolvendo para λ, resulta

$$\lambda = \frac{\boldsymbol{\xi}^T \mathbf{V} \boldsymbol{\xi}}{\boldsymbol{\xi}^T \mathbf{T} \boldsymbol{\xi}} > 0 \tag{5.6.4}$$

porque tanto o numerador quanto o denominador são números reais positivos em virtude das Eqs.(5.3.7) e (5.6.1).

Ortogonalidade dos Vetores Característicos

No Exemplo 5.5.1 os modos de freqüências ω_1 e ω_2 eram automaticamente ortogonais no produto interno (5.5.5). Isto não é uma ocorrência fortuita, mas tem lugar sempre que as freqüências são *distintas*. De fato, sejam $\boldsymbol{\varrho}^{(1)}$ e $\boldsymbol{\varrho}^{(2)}$ soluções de (5.6.3) com diferentes valores de λ:

$$\mathbf{V} \boldsymbol{\varrho}^{(1)} = \lambda_1 \mathbf{T} \boldsymbol{\varrho}^{(1)} \; ; \tag{5.6.5a}$$

$$\mathbf{V} \boldsymbol{\varrho}^{(2)} = \lambda_2 \mathbf{T} \boldsymbol{\varrho}^{(2)} \; . \tag{5.6.5b}$$

Multiplicando (5.6.5a) pela esquerda por ${\boldsymbol{\varrho}^{(2)}}^T$, (5.6.5b) por ${\boldsymbol{\varrho}^{(1)}}^T$ e subtraindo as equações resultantes obtém-se

$$(\lambda_1 - \lambda_2)(\boldsymbol{\varrho}^{(1)}, \boldsymbol{\varrho}^{(2)}) = {\boldsymbol{\varrho}^{(2)}}^T \mathbf{V} \boldsymbol{\varrho}^{(1)} - {\boldsymbol{\varrho}^{(1)}}^T \mathbf{V} \boldsymbol{\varrho}^{(2)} \; , \tag{5.6.6}$$

onde usamos a simetria do produto interno (5.5.5). Mas, como $\mathbf{V}$ é também matriz simétrica, o Exercício 5.6.1 estabelece que

$${\boldsymbol{\varrho}^{(2)}}^T \mathbf{V} \boldsymbol{\varrho}^{(1)} = {\boldsymbol{\varrho}^{(1)}}^T \mathbf{V} \boldsymbol{\varrho}^{(2)} \; . \tag{5.6.7}$$

Conseqüentemente,

$$(\lambda_1 - \lambda_2)\,(\boldsymbol{\varrho}^{(1)}, \boldsymbol{\varrho}^{(2)}) = 0 \quad \Longrightarrow \quad (\boldsymbol{\varrho}^{(1)}, \boldsymbol{\varrho}^{(2)}) = 0 \tag{5.6.8}$$

se $\lambda_1 \neq \lambda_2$. Se todas as freqüências forem distintas, as soluções linearmente independentes de (5.4.5) serão automaticamente ortogonais entre si, bastando normalizá-las no produto interno (5.5.5) para se construir a matriz modal. Se o espectro de freqüências for degenerado, isto é, se a equação característica (5.4.6) possuir raízes múltiplas, será necessário escolher combinações lineares convenientes das soluções linearmente independentes de (5.4.5) associadas a uma mesma freqüência característica, de modo a se obter um conjunto de vetores mutuamente ortogonais.

Completeza dos Modos Normais de Vibração

Finalmente, outra suposição feita para permitir escrever a solução geral das equações de movimento na forma (5.4.13) foi a da existência de exatamente n soluções linearmente independentes $\boldsymbol{\varrho}^{(1)}, \ldots, \boldsymbol{\varrho}^{(n)}$ da Eq.(5.4.5). A fim de demonstrar isto, notemos inicialmente que a positividade da matriz $\mathbf{T}$ assegura a existência de sua inversa $\mathbf{T}^{-1}$. Com

efeito, se $\mathbf{T}$ não tivesse inversa existiria um vetor $\boldsymbol{\xi} \neq 0$ tal que $\mathbf{T}\boldsymbol{\xi} = 0$. Multiplicando esta última equação pela esquerda por $\boldsymbol{\xi}^T$ seríamos levados a concluir que $(\boldsymbol{\xi}, \boldsymbol{\xi}) = 0$ donde $\boldsymbol{\xi} = 0$, o que vem a ser uma contradição. Escrevamos, portanto, a Eq.(5.4.5) na forma

$$\mathbf{T}^{-1}\mathbf{V}\boldsymbol{\xi} = \lambda\boldsymbol{\xi} \ , \tag{5.6.9}$$

que é a equação de autovalores e autovetores para o operador $\mathbf{K} = \mathbf{T}^{-1}\mathbf{V}$. De acordo com o Teorema C.13, demonstrado no Apêndice C, se provarmos que $\mathbf{K}$ é simétrico estará provado que $\mathbf{K}$ é auto-adjunto e assegurada a existência de uma base ortonormal de $\boldsymbol{R}^n$ (e, portanto, de uma coleção de n vetores linearmente independentes) constituída por autovetores de $\mathbf{K}$. Por outro lado, para provar que um operador é simétrico basta verificar que sua matriz associada numa base ortonormal arbitrária $\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(n)}$ é simétrica. Mas, por definição, os elementos de matriz de $\mathbf{K}$ são

$$K_{ij} = (\boldsymbol{\xi}^{(i)}, \mathbf{K}\boldsymbol{\xi}^{(j)}) = \boldsymbol{\xi}^{(i)^T}\mathbf{T}\mathbf{T}^{-1}\mathbf{V}\boldsymbol{\xi}^{(j)} = \boldsymbol{\xi}^{(i)^T}\mathbf{V}\boldsymbol{\xi}^{(j)} \ , \tag{5.6.10}$$

donde, usando a Eq.(5.6.7), infere-se imediatamente que $K_{ij} = K_{ji}$. Fica, portanto, estabelecida a asserção feita na Seção 5.4 de que é possível encontrar n soluções linearmente independentes da Eq.(5.4.5).

Como último comentário, note que a positividade da matriz $\mathbf{V}$ não foi utilizada na demonstração de que a matriz modal diagonaliza simultaneamente $\mathbf{T}$ e $\mathbf{V}$. Assim, de um ponto de vista estritamente matemático, provamos que, dadas duas formas quadráticas reais, basta que uma delas seja positiva definida para que exista uma matriz capaz de diagonalizá-las simultaneamente. Essa diagonalização simultânea é geometricamente visualizável sem grande dificuldade no caso de sistemas com apenas dois graus de liberdade (Aravind 1989).

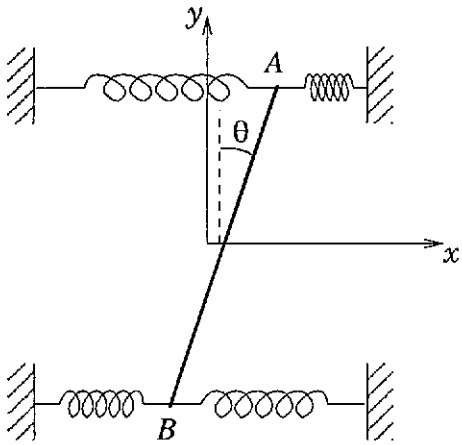


Figura 5.9: Problema 5.5.

PROBLEMAS

5.1. Uma partícula de massa m move-se no semi-eixo $x > 0$ com energia potencial

$$V(x) = V_0 \left(e^{-kx} + Cx\right) ,$$

onde todas as constantes são positivas. Determine para que valores das constantes existem posições de equilíbrio e encontre a freqüência das pequenas oscilações em torno das posições de equilíbrio estável.

5.2. Uma conta de massa m desliza sem atrito ao longo de uma circunferência de raio R que gira em torno de seu diâmetro vertical com velocidade angular constante ω. Sendo a origem tomada no centro da circunferência e θ o ângulo definido pela partícula relativamente à direção vertical, as coordenadas cartesianas da partícula de massa m são $x = R\,\mathrm{sen}\,\theta\cos\varphi$, $y = R\,\mathrm{sen}\,\theta\,\mathrm{sen}\,\varphi$, $z = -R\cos\theta$, onde $\varphi = \omega t$ é o ângulo de rotação em torno do eixo vertical z. (i) Se o nível zero do potencial gravitacional for tomado no ponto mais baixo da circunferência, mostre que a lagrangiana do sistema é dada por

$$L = \frac{m}{2}R^2\dot{\theta}^2 - mgR\left[(1-\cos\theta) - \frac{1}{2}\frac{\omega^2}{\omega_c^2}\,\mathrm{sen}^2\theta\right] ,$$

onde $\omega_c = \sqrt{g/R}$. (ii) Estude as configurações de equilíbrio e classifique-as quanto à estabilidade. (iii) Mostre que este sistema também apresenta quebra espontânea de simetria (Sivardière 1983).

5.3. Um pêndulo duplo genérico tem a lagrangiana dada no Exemplo 1.5.3. Considere pequenas oscilações no caso particular $l_1 = l_2 = l$, $m_1 = m$, $m_2 = m/3$. (i) Determine os modos normais de vibração e represente-os graficamente. (ii) Encontre a posição de cada partícula no instante t se em $t = 0$ ambas estavam em repouso deslocadas do mesmo ângulo pequeno θ_0 em relação à vertical. (iii) Obtenha as coordenadas normais do sistema.

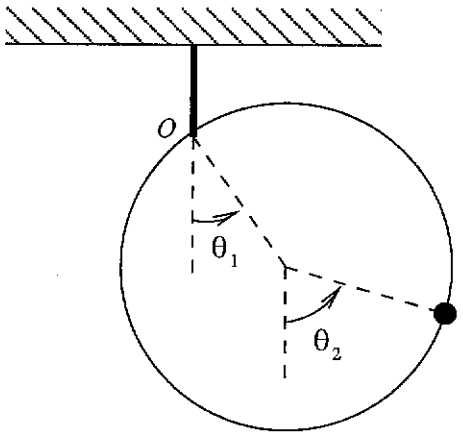


Figura 5.10: Problema 5.7.

5.4. Uma partícula move-se sem atrito na parte interna côncava de uma superfície de revolução $z = f(\rho)$, onde ρ, θ, z são coordenadas cilíndricas, e na presença de um campo gravitacional constante $\mathbf{g} = -g\hat{\mathbf{k}}$. (i) Encontre a condição para que exista uma órbita circular de raio ρ_0. (ii) Mostre que uma órbita circular é estável ou instável conforme $3f'(\rho_0) + \rho_0 f''(\rho_0)$ seja positivo ou negativo. (iii) Para as órbitas estáveis, encontre a freqüência Ω das pequenas oscilações em torno da configuração de equilíbrio.

5.5. O sistema representado na figura está sobre uma mesa horizontal lisa. Todas as molas são iguais e a haste rígida homogênea AB tem massa m e comprimento $2l$. Não há movimento da haste na direção y. Usando como coordenadas a abscissa x do centro de massa da haste e o ângulo pequeno θ que ela forma com o eixo y, determine: (i) a lagrangiana; (ii) as freqüências características e os modos normais de vibração. (iii) Represente graficamente os modos normais.

5.6. Duas partículas de mesma massa m estão ligadas entre si por uma mola de constante elástica k' e por molas de mesma constante elástica k a dois suportes fixos. Na posição de equilíbrio as molas têm todas o mesmo comprimento natural e o movimento está restrito a uma linha reta. (i) Determine as freqüências características e os modos normais de vibração do sistema. (ii) Justifique fisicamente o valor da menor freqüência característica. (iii) Represente graficamente os modos normais de vibração. (iv) Obtenha a matriz modal e as coordenadas normais. Verifique explicitamente que as coordenadas normais diagonalizam a lagrangiana. (v) Encontre a posição de cada partícula no instante t dadas as seguintes condições iniciais em $t = 0$: (a) ambas as partículas na posição de equilíbrio, a primeira com velocidade v_0 e a segunda em repouso; (b) ambas as partículas em repouso, a primeira na posição de equilíbrio e a segunda deslocada de a. No caso (b), descreva qualitativamente o movimento de cada partícula se $k' \ll k$.

5.7. Um aro fino de massa m e raio R oscila num plano vertical com um ponto fixo O (ver figura). Ao longo do aro desliza sem atrito uma conta com a mesma massa m. (i) Mostre que a lagrangiana do sistema é

$$L = \frac{3mR^2}{2}\,\dot{\theta}_1^2 + \frac{mR^2}{2}\,\dot{\theta}_2^2 + mR^2\,\dot{\theta}_1\,\dot{\theta}_2\cos(\theta_1 - \theta_2) + 2mgR\cos\theta_1 + mgR\cos\theta_2 \ .$$

(ii) Considerando pequenas oscilações, encontre as freqüências características e os modos normais de vibração. Represente graficamente os modos normais. (iii) Obtenha a solução das equações de movimento com as condições iniciais $\theta_1(0) = 0$, $\theta_2(0) = \theta_0$, $\dot{\theta}_1(0) = \dot{\theta}_2(0) = 0$. (iv) Determine a matriz modal e as coordenadas normais do sistema.

5.8. Dada a lagrangiana $L(q, \dot{q}, t)$ considere a nova lagrangiana $\lambda(q, \dot{q}, \eta, \dot{\eta}, t)$ definida por (Núñes-Yepes, Delgado & Salas-Brito 2001)

$$\lambda(q, \dot{q}, \eta, \dot{\eta}, t) = \sum_{l=1}^{n} \left(\frac{\partial L}{\partial q_l} \eta_l + \frac{\partial L}{\partial \dot{q}_l} \dot{\eta}_l \right) .$$

(i) Mostre que o princípio variacional $\delta \int \lambda(q, \dot{q}, \eta, \dot{\eta}, t) dt = 0$ fornece de um só golpe as equações de Lagrange para $q_k(t)$ e as equações de movimento para os pequenos desvios $\eta_k(t)$ entre duas trajetórias dinâmicas $q_k(t)$ e $\bar{q}_k(t) = q_k(t) + \eta_k(t)$, isto é, prove que as equações

$$\frac{d}{dt}\left(\frac{\partial \lambda}{\partial \dot{\eta}_k}\right) - \frac{\partial \lambda}{\partial \eta_k} = 0 \quad , \quad \frac{d}{dt}\left(\frac{\partial \lambda}{\partial \dot{q}_k}\right) - \frac{\partial \lambda}{\partial q_k} = 0$$

equivalem a

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = 0 \quad , \quad \sum_{l=1}^{n} \left(M_{kl} \ddot{\eta}_l + C_{kl} \dot{\eta}_l + K_{kl} \eta_l \right) = 0$$

onde

$$M_{kl} = \frac{\partial^2 L}{\partial \dot{q}_k \partial \dot{q}_l} \ , \ C_{kl} = \frac{d}{dt}\left(\frac{\partial^2 L}{\partial \dot{q}_k \partial \dot{q}_l}\right) + \frac{\partial^2 L}{\partial \dot{q}_k \partial q_l} - \frac{\partial^2 L}{\partial q_k \partial \dot{q}_l} \ , \ K_{kl} = \frac{d}{dt}\left(\frac{\partial^2 L}{\partial \dot{q}_k \partial q_l}\right) - \frac{\partial^2 L}{\partial q_k \partial q_l} \ .$$

(ii) Para a Lagrangiana $L = \frac{1}{2} \sum_{k,l=1}^{n} m_{kl} \dot{q}_k \dot{q}_l - V(q)$ com os coeficientes constantes $m_{kl} = m_{lk}$, mostre que

$$\sum_{l=1}^{n} \left(m_{kl} \ddot{\eta}_l + \frac{\partial^2 V}{\partial q_k \partial q_l} \eta_l \right) = 0$$

e comprove diretamente que estas equações coincidem com as que são satisfeitas pelos pequenos desvios entre duas soluções de

$$\sum_{l=1}^{n} m_{kl} \ddot{q}_l = -\frac{\partial V}{\partial q_k} \ .$$

5.9. A lagrangiana

$$L = \frac{m}{2} \sum_{l=1}^{N-1} \dot{\eta}_l^2 - \frac{k}{2} \sum_{l=0}^{N-1} (\eta_{l+1} - \eta_l)^2$$

descreve uma cadeia linear de $N-1$ osciladores ($N \geq 2$) de mesma massa m conectados por molas iguais com constante de força k. Os extremos da cadeia são fixos, isto é,

$$\eta_0 = 0 \quad , \quad \eta_N = 0 \ .$$

(i) Encontre as equações de movimento para $\eta_1, \cdots, \eta_{N-1}$ e mostre que elas admitem os modos normais de vibração

$$\eta_l^{(s)} = C^{(s)} \operatorname{sen} \frac{sl\pi}{N} \cos(\omega_s t + \phi_s) \equiv \varrho_l^{(s)} \cos(\omega_s t + \phi_s) \quad , \quad s, l = 1, \cdots, N-1$$

com as freqüências características

$$\omega_s = 2\sqrt{\frac{k}{m}} \operatorname{sen} \frac{s\pi}{2N} \ .$$

(ii) Como todas as freqüências são distintas (convença-se disto!), os vetores $\{\boldsymbol{\varrho}^{(s)}\}$ são mutuamente ortogonais. Mostre que $\boldsymbol{\varrho}^{(s)}$ é normalizado pela escolha $C^{(s)} = (2/Nm)^{1/2}$, a mesma para todos os vetores. Sugestão: para provar que $\sum_{l=1}^{N-1} \operatorname{sen}^2(ls\pi/N) = N/2$, escreva $\operatorname{sen}^2\theta = (1 - \cos 2\theta)/2$, use $\cos\theta = \operatorname{Re} e^{i\theta}$ para θ real, e note que $\sum_{l=1}^{N-1} e^{ilx}$ é a soma de uma progressão geométrica. (iii) Se no instante $t = 0$ só o primeiro oscilador está deslocado de sua posição de equilíbrio e todos os osciladores estão em repouso, determine $\eta_l(t)$. Sugestão: explore a ortogonalidade dos vetores $\{\boldsymbol{\varrho}^{(s)}\}$. (iv) Um bônus puramente matemático deste problema é a curiosa identidade trigonométrica

$$\sum_{j=1}^{N-1} \operatorname{sen}\left(j\frac{r\pi}{N}\right) \operatorname{sen}\left(j\frac{s\pi}{N}\right) = \frac{N}{2}\delta_{rs} \ , \quad r, s = 1, \ldots, N-1 \ , \ N \geq 2 \ .$$

Prove que esta identidade decorre imediatamente das relações que caracterizam os vetores $\{\boldsymbol{\varrho}^{(s)}\}$, com $C^{(s)} = \sqrt{2/Nm}$, como um conjunto ortonormal.

Capítulo 6

MECÂNICA RELATIVÍSTICA

> I am convinced that the philosophers have had a harmful effect upon the progress of scientific thinking in removing certain fundamental concepts from the domain of empiricism... This is particularly true of our concepts of time and space, which physicists have been obliged by the facts to bring down from the Olympus of the *a priori* in order to adjust them and put them in a serviceable condition.
>
> Albert Einstein

Quando as partículas movem-se com velocidades próximas à velocidade da luz no vácuo, a mecânica newtoniana entra em contradição com a experiência e precisa ser reformulada à luz da teoria especial da relatividade. Suporemos previamente conhecidas as bases físicas da teoria special da relatividade (Bergmann 1975; Rindler 1982) bem como as transformações de Lorentz e suas conseqüências mais imediatas.[1] Nossa atenção estará voltada para o desenvolvimento de um formalismo que permita formular as leis da Física numa forma manifestamente válida em todos os sistemas de referência inerciais, seguido de sua aplicação à dinâmica de uma partícula e sua incorporação ao arcabouço da mecânica analítica.

6.1 Transformações de Lorentz

Seja $\mathcal{K}$ um referencial inercial e $\mathcal{K}'$ um outro referencial inercial que se move em relação a $\mathcal{K}$ com velocidade constante v, conforme a Fig. 6.1 (os eixos de $\mathcal{K}'$ são paralelos aos eixos de $\mathcal{K}$). A fim de evitar complicações desnecessárias, suponhamos que as origens O e O' coincidam nos instantes $t = t' = 0$ e que a velocidade relativa v seja paralela ao eixo x de $\mathcal{K}$ (esta última restrição será relaxada mais adiante). Então as coordenadas (x, y, z, t) e (x', y', z', t'), atribuídas a um mesmo *evento* por observadores fixos nos respectivos referenciais, estão relacionadas da seguinte maneira:

$$x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}} \ , \tag{6.1.1a}$$

$$y' = y \ , \tag{6.1.1b}$$

$$z' = z \ , \tag{6.1.1c}$$

[1] Os principais trabalhos originais podem ser encontrados em Lorentz, Einstein, Minkowski & Weyl (1952).

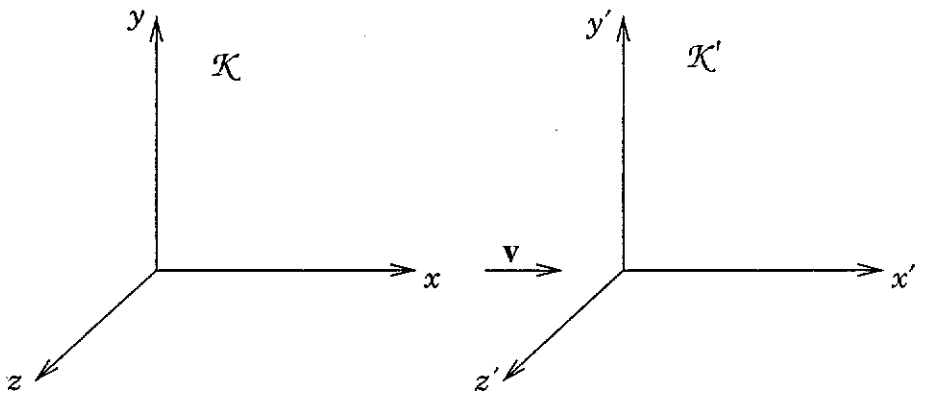


Figura 6.1: Transformação de Lorentz.

$$t' = \frac{t - vx/c^2}{\sqrt{1 - v^2/c^2}} \ . \qquad (6.1.1d)$$

As equações (6.1.1) constituem uma transformação de Lorentz e foram deduzidas por Einstein em 1905 com base no postulado fundamental da constância da velocidade da luz no vácuo.[2] Formalmente, o limite não relativístico da transformação de Lorentz obtém-se fazendo $c \to \infty$. De fato, passando ao limite $c \to \infty$ a Eq.(6.1.1) reduz-se a

$$x' = x - vt \quad , \quad y' = y \quad , \quad z' = z \quad , \quad t' = t \ , \qquad (6.1.2)$$

que é uma transformação de Galileu, característica da mecânica newtoniana.

É fácil e conveniente generalizar (6.1.1) de modo a eliminar a particularidade da escolha da velocidade $\mathbf{v}$ paralela à direção x. Em termos de $\mathbf{v} = v\hat{\mathbf{i}}$ é claro que

$$vx = \mathbf{v} \cdot \mathbf{r} \ , \qquad (6.1.3)$$

onde $\mathbf{r} = (x, y, z)$ é o vetor posição referente a $\mathcal{K}$. Por outro lado, definindo

$$\gamma = (1 - v^2/c^2)^{-1/2} \quad , \quad \boldsymbol{\beta} = \mathbf{v}/c \ , \qquad (6.1.4)$$

podemos escrever a parte espacial de (6.1.1) na forma vetorial

$$\mathbf{r}' = \mathbf{r} + (\gamma - 1)x\,\hat{\mathbf{i}} - \gamma \mathbf{v}\, t = \mathbf{r} + (\gamma - 1)\frac{\mathbf{v} \cdot \mathbf{r}}{v}\frac{\mathbf{v}}{v} - \gamma \mathbf{v}\, t \ , \qquad (6.1.5)$$

onde usamos (6.1.3). Finalmente, as Eqs.(6.1.1) podem ser postas na forma

$$\mathbf{r}' = \mathbf{r} + (\gamma - 1)\frac{\boldsymbol{\beta} \cdot \mathbf{r}}{\beta^2}\boldsymbol{\beta} - \gamma c \boldsymbol{\beta} t \ , \qquad (6.1.6a)$$

[2] Embora a transformação de Lorentz (6.1.1) já fosse conhecida e empregada por Lorentz antes de 1905, coube a Einstein deduzi-la e interpretá-la de forma revolucionária.

$$t' = \gamma\left(t - \frac{\boldsymbol{\beta}\cdot \mathrm{r}}{c}\right) \ , \tag{6.1.6b}$$

que não faz menção a nenhuma direção privilegiada para $\mathbf{v} = c\,\boldsymbol{\beta}$.

Intervalo Invariante

O *intervalo* ds entre dois eventos infinitesimalmente próximos, com coordenadas (x, y, z, t) e $(x + dx, y + dy, z + dz, t + dt)$, respectivamente, é definido por[3]

$$ds^2 = c^2\,dt^2 - dx^2 - dy^2 - dz^2 \ . \tag{6.1.7}$$

Das Eqs.(6.1.1) deduzimos

$$dx' = \gamma\,(dx - v dt) \quad , \quad dy' = dy \quad , \quad dz' = dz \quad , \quad dt' = \gamma\,(dt - v dx/c^2) \ , \tag{6.1.8}$$

e um cálculo algébrico simples mostra que

$$ds'^2 = ds^2 \ . \tag{6.1.9}$$

Este é um importante resultado: o intervalo entre dois eventos é invariante sob transformações de Lorentz. Diz-se, também, que ds é um *escalar de Lorentz*. Naturalmente, o mesmo resultado vale para um intervalo finito Δs. Note que ds é análogo à distância entre dois pontos infinitesimalmente próximos num espaço euclidiano tridimensional. No entanto, no caso da teoria especial da relatividade nem a distância espacial nem o intervalo de tempo são separadamente invariantes, mas somente a particular combinação de espaço e tempo expressa em (6.1.7).[4] O espaço-tempo quadridimensional com "distância" invariante ds definida por (6.1.7) é conhecido com *espaço de Minkowski*.

De agora em diante faremos uso da seguinte notação indicial:

$$x^0 = c\,t \quad , \quad x^1 = x \quad , \quad x^2 = y \quad , \quad x^3 = z \ . \tag{6.1.10}$$

A transformação de Lorentz (6.1.6) é uma transformação linear que pode ser escrita na forma

$$x'^{\mu} = \Lambda^{\mu}{}_0\, x^0 + \Lambda^{\mu}{}_1\, x^1 + \Lambda^{\mu}{}_2\, x^2 + \Lambda^{\mu}{}_3\, x^3 \equiv \Lambda^{\mu}{}_{\nu}\, x^{\nu} \ , \tag{6.1.11}$$

onde passamos a utilizar a convenção de soma sobre índices repetidos de Einstein: toda vez que um índice aparecer repetido, tal como ν nesta última equação, deve-se efetuar

[3] Esta definição não é universal. Muitos autores preferem escrever $ds^2 = -c^2\,dt^2 + dx^2 + dy^2 + dz^2$, de modo que é preciso atentar para as convenções utilizadas antes de comparar fórmulas de diferentes autores.

[4] Daí a célebre e dramática frase de Minkowski ao apresentar essas idéias pela primeira vez em 1908: "De ora em diante, o espaço por si mesmo e o tempo por si mesmo estão fadados a desvanecer-se em meras sombras, e apenas uma espécie de união dos dois preservará uma realidade independente."

uma soma de 0 a 3 sobre tal índice. Convencionaremos, também, que índices gregos tais como $\alpha, \beta, \mu, \nu, \ldots$ assumem os valores $0, 1, 2, 3$, ao passo que índices latinos tais como $i, j, k, l, \ldots$ tomam os valores $1, 2, 3$. Os coeficientes da transformação de Lorentz $\Lambda^{\mu}{}_{\nu}$ formam uma matriz $\mathbf{\Lambda} = (\Lambda^{\mu}{}_{\nu})$ chamada de matriz da transformação de Lorentz.[5] No caso da transformação de Lorentz ao longo da direção x, a matriz associada é

$$\mathbf{\Lambda} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{6.1.12}$$

■ **Exercício 6.1.1.** Prove que os elementos da matriz $\mathbf{\Lambda}$ são

$$\Lambda^{0}{}_{0} = \gamma \quad , \quad \Lambda^{0}{}_{i} = \Lambda^{i}{}_{0} = -\gamma\beta_i \ , \tag{6.1.13a}$$

$$\Lambda^{i}{}_{j} = \delta_{ij} + (\gamma - 1)\frac{\beta_i \beta_j}{\beta^2} \ , \tag{6.1.13b}$$

no caso da transformação de Lorentz (6.1.6), em que a direção de $\mathbf{v} = c\boldsymbol{\beta}$ é arbitrária e, naturalmente, $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3) \equiv (\beta_x, \beta_y, \beta_z)$. ■

Tensor Métrico

A introdução da matriz $\mathbf{G}$, com elementos $g_{\mu\nu}$, definida por[6]

$$\mathbf{G} = (g_{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{6.1.14}$$

permite escrever

$$ds^2 = c^2\,dt^2 - dx^2 - dy^2 - dz^2 \equiv g_{\mu\nu}\,dx^{\mu}\,dx^{\nu} \ , \tag{6.1.15}$$

nesta última equação havendo, conforme convencionado, uma soma dupla nos índices repetidos μ e ν. A matriz $\mathbf{G}$ define uma *métrica* no espaço de Minkowski e seus elementos $g_{\mu\nu}$ são chamados de componentes do *tensor métrico*.

Grupo de Lorentz

A transformação de Lorentz mais geral possível é definida como qualquer transformação linear da forma (6.1.11) que deixa ds^2 invariante, isto é,

$$ds^2 = g_{\mu\nu}\,dx^{\mu}\,dx^{\nu} = ds'^2 = g_{\mu\nu}\,dx'^{\mu}\,dx'^{\nu} = g_{\mu\nu}\,\Lambda^{\mu}{}_{\alpha}\,dx^{\alpha}\,\Lambda^{\nu}{}_{\beta}\,dx^{\beta} \ , \tag{6.1.16}$$

[5] Em $\Lambda^{\mu}{}_{\nu}$ o índice de linha é μ e o índice de coluna é ν.

[6] Diversos autores preferem definir $\mathbf{G}$ com os sinais de todos os elementos da diagonal trocados.

para deslocamentos dxs arbitrários. A fim de poder comparar os coeficientes dos deslocamentos no segundo e no último membros de (6.1.16), note que (vide Apêndice A)

$$g_{\mu\nu}\, dx^\mu\, dx^\nu \equiv g_{\alpha\beta}\, dx^\alpha\, dx^\beta \ . \tag{6.1.17}$$

Assim, resulta

$$\Lambda^\mu{}_\alpha\, g_{\mu\nu}\, \Lambda^\nu{}_\beta\, dx^\alpha\, dx^\beta = g_{\alpha\beta}\, dx^\alpha\, dx^\beta \ , \tag{6.1.18}$$

donde, face à arbitrariedade dos dxs,

$$\Lambda^\mu{}_\alpha\, g_{\mu\nu}\, \Lambda^\nu{}_\beta = g_{\alpha\beta} \ . \tag{6.1.19}$$

Esta é a equação que caracteriza $\boldsymbol{\Lambda}$ como matriz de uma transformação de Lorentz. Em notação matricial a Eq.(6.1.19) equivale a

$$\boldsymbol{\Lambda}^T \mathbf{G}\, \boldsymbol{\Lambda} = \mathbf{G} \ . \tag{6.1.20}$$

Multiplicando esta última equação pela direita por $\mathbf{G}$ e notando que $\mathbf{G}^2 = \boldsymbol{I}$, resulta $\boldsymbol{\Lambda}^T \mathbf{G} \boldsymbol{\Lambda} \mathbf{G} = \boldsymbol{I}$, o que mostra que $\boldsymbol{\Lambda}^T \mathbf{G}$ é a matriz inversa de $\boldsymbol{\Lambda}\mathbf{G}$. Como a inversa de uma matriz pela esquerda coincide com a inversa pela direita, temos também $\boldsymbol{\Lambda}\mathbf{G}\boldsymbol{\Lambda}^T\mathbf{G} = \boldsymbol{I}$, donde

$$\boldsymbol{\Lambda}\mathbf{G}\boldsymbol{\Lambda}^T = \mathbf{G} \ , \tag{6.1.21}$$

que é uma condição equivalente a (6.1.20). Uma transformação de Lorentz é uma espécie de rotação no espaço-tempo, e a matriz associada é ortogonal em relação à métrica $\mathbf{G}$ do espaço de Minkowski. No caso das rotações tridimensionais a métrica é dada pela matriz identidade, e a condição correspondente a (6.1.20) é a Eq.(3.1.11).

■ **Exercício 6.1.2.** Usando (6.1.20) e (6.1.21), prove que o conjunto das transformações de Lorentz forma um grupo, chamado de grupo de Lorentz. Sugestões: (i) a definição de grupo está no final da Seção 3.1; (ii) se $\mathbf{G}$ e $\mathbf{H}$ são matrizes inversíveis então $(\mathbf{GH})^{-1} = \mathbf{H}^{-1}\mathbf{G}^{-1}$; (iii) a transposta da inversa é igual à inversa da transposta. ■

Vale ressaltar que a definição (6.1.20) de transformação de Lorentz abrange uma classe muito mais ampla de transformações lineares do que as transformações de Lorentz puras da forma (6.1.6). Por exemplo, uma rotação dos eixos espaciais constitui uma transformação de Lorentz consoante nossa definição geral, pois não afeta o tempo e deixa invariante a distância espacial entre dois pontos. A inversão espacial $t' = t$, $\mathbf{r}' = -\mathbf{r}$ e a inversão temporal $t' = -t$, $\mathbf{r}' = \mathbf{r}$ também constituem exemplos de transformações de Lorentz, pois ambas preservam ds^2.

As principais propriedades das transformações de Lorentz podem ser inferidas de (6.1.19) ou (6.1.20). Tomando o determinante de (6.1.20) deduzimos

$$-1 = \det\mathbf{G} = \det\left(\boldsymbol{\Lambda}^T \mathbf{G}\, \boldsymbol{\Lambda}\right) = \det\boldsymbol{\Lambda}^T \det\mathbf{G} \det\boldsymbol{\Lambda} = (\det\boldsymbol{\Lambda})\,(-1)\,(\det\boldsymbol{\Lambda}) \ , \tag{6.1.22}$$

isto é,

$$(\det \mathbf{\Lambda})^2 = 1 \quad \Longrightarrow \quad \det \mathbf{\Lambda} = \pm 1 \ . \tag{6.1.23}$$

Uma transformação de Lorentz com determinante $+1$ é dita *própria*, e é dita *imprópria* se o determinante é -1. Qualquer transformação de Lorentz pura, tal como (6.1.1), é própria (verifique!).

Tomando $\alpha = \beta = 0$ em (6.1.19) obtém-se

$$1 = g_{00} = \Lambda^{\mu}{}_0 \, g_{\mu\nu} \, \Lambda^{\nu}{}_0 = (\Lambda^0{}_0)^2 - (\Lambda^1{}_0)^2 - (\Lambda^2{}_0)^2 - (\Lambda^3{}_0)^2 \ , \tag{6.1.24}$$

ou seja,

$$(\Lambda^0{}_0)^2 = 1 + (\Lambda^1{}_0)^2 + (\Lambda^2{}_0)^2 + (\Lambda^3{}_0)^2 \geq 1 \ , \tag{6.1.25}$$

donde, finalmente,

$$\Lambda^0{}_0 \geq +1 \quad \text{ou} \quad \Lambda^0{}_0 \leq -1 \ . \tag{6.1.26}$$

Se $\Lambda^0{}_0 \geq +1$ a transformação de Lorentz é dita ortócrona (preserva o sentido do tempo); se $\Lambda^0{}_0 \leq -1$ a transformação de Lorentz é dita não-ortócrona (inverte o sentido do tempo). Qualquer transformação de Lorentz pura, que conecta referenciais com eixos paralelos e em movimento relativo com velocidade constante, é ortócrona.[7] Uma inversão temporal é uma transformação de Lorentz não-ortócrona. As transformações próprias e ortócronas são chamadas de transformações de Lorentz *restritas*. As transformações de Lorentz infinitesimais desempenham um papel muito importante no estudo das propriedades do grupo de Lorentz. Uma transformação de Lorentz infinitesimal é necessariamente restrita porque só difere infinitesimalmente da transformação identidade, que é própria e ortócrona, e uma variação infinitesimal não pode provocar um salto finito nem em $\Lambda^0{}_0$ nem no determinante de $\mathbf{\Lambda}$. Portanto, somente as transformações de Lorentz restritas podem ser construídas por aplicação de sucessivas transformações de Lorentz infinitesimais. A presente discussão estabelece que o conjunto das transformações de Lorentz se decompõe em quatro subconjuntos disjuntos, chamados de setores do grupo de Lorentz e denotados por $L_+^{\uparrow}, L_+^{\downarrow}, L_-^{\uparrow}$ e $L_-^{\downarrow}$.

Setores do Grupo de Lorentz

$$\begin{aligned}
L_+^{\uparrow} &: \det \mathbf{\Lambda} = +1 \quad \text{e} \quad \Lambda^0{}_0 \geq +1 \\
L_+^{\downarrow} &: \det \mathbf{\Lambda} = +1 \quad \text{e} \quad \Lambda^0{}_0 \leq -1 \\
L_-^{\uparrow} &: \det \mathbf{\Lambda} = -1 \quad \text{e} \quad \Lambda^0{}_0 \geq +1 \\
L_-^{\downarrow} &: \det \mathbf{\Lambda} = -1 \quad \text{e} \quad \Lambda^0{}_0 \leq -1
\end{aligned}$$

O conjunto da transformações de Lorentz restritas $L_+^{\uparrow}$ é o único setor que forma um subgrupo do grupo de Lorentz. Este subgrupo é chamado de *grupo de Lorentz restrito*.

A única parte não trivial da demonstração de que $L_+^{\uparrow}$ constitui um grupo é a verificação de que o produto de duas matrizes de Lorentz ortócronas é também uma matriz ortócrona.

[7] As transformações de Lorentz puras também costumam ser chamadas de "boosts".

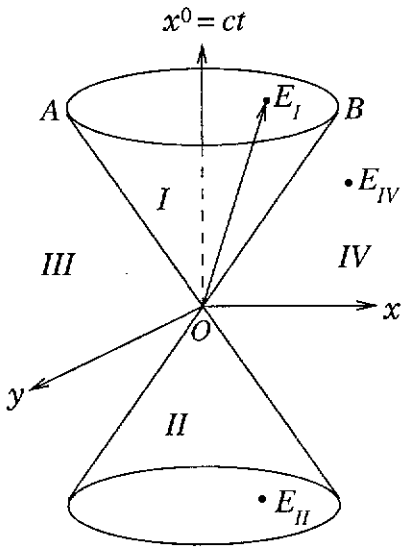


Figura 6.2: Cone de luz do evento O (uma dimensão espacial foi suprimida).

Sejam $\mathbf{\Lambda}$ e $\bar{\mathbf{\Lambda}}$ duas transformações de Lorentz ortócronas consecutivas e seja $\bar{\bar{\mathbf{\Lambda}}} = \bar{\mathbf{\Lambda}}\mathbf{\Lambda}$. Então $\bar{\bar{\Lambda}}^{\mu}{}_{\nu} = \bar{\Lambda}^{\mu}{}_{\alpha}\Lambda^{\alpha}{}_{\nu}$, de modo que

$$\bar{\bar{\Lambda}}^{0}{}_{0} = \bar{\Lambda}^{0}{}_{0}\Lambda^{0}{}_{0} + \bar{\Lambda}^{0}{}_{1}\Lambda^{1}{}_{0} + \bar{\Lambda}^{0}{}_{2}\Lambda^{2}{}_{0} + \bar{\Lambda}^{0}{}_{3}\Lambda^{3}{}_{0} \ . \tag{6.1.27}$$

Pondo $\mathbf{a} = (\bar{\Lambda}^{0}{}_{1}, \bar{\Lambda}^{0}{}_{2}, \bar{\Lambda}^{0}{}_{3})$, $\mathbf{b} = (\Lambda^{1}{}_{0}, \Lambda^{2}{}_{0}, \Lambda^{3}{}_{0})$ e usando (6.1.25) bem como sua análoga decorrente de (6.1.21), ficamos com

$$\bar{\bar{\Lambda}}^{0}{}_{0} = \sqrt{1+a^2}\,\sqrt{1+b^2} + \mathbf{a}\cdot\mathbf{b} \ , \tag{6.1.28}$$

onde $a = |\mathbf{a}|$ e $b = |\mathbf{b}|$. Mas

$$\sqrt{1+a^2}\sqrt{1+b^2} = \sqrt{1+a^2+b^2+a^2b^2} = \sqrt{(1+ab)^2+(a-b)^2} \geq 1+ab \ , \tag{6.1.29}$$

donde

$$\bar{\bar{\Lambda}}^{0}{}_{0} \geq 1 + ab + \mathbf{a}\cdot\mathbf{b} \ . \tag{6.1.30}$$

Levando em conta que $\mathbf{a}\cdot\mathbf{b} \geq -ab$, obtemos $\bar{\bar{\Lambda}}^{0}{}_{0} \geq 1$, como queríamos demonstrar.

6.2 Cone de Luz e Causalidade

Um intervalo $\Delta s^2 = c^2\,\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2$ entre dois eventos é dito *do tipo espaço* se a parte espacial é maior do que a parte temporal ($\Delta s^2 < 0$), e *do tipo tempo* se a parte temporal é maior do que a parte espacial ($\Delta s^2 > 0$). Se $\Delta s^2 = 0$ o intervalo é dito *do tipo luz* . Note que esta classificação dos intervalos é absoluta, em outras palavras, é válida em todos os referenciais inerciais porque Δs^2 é um escalar (invariante).

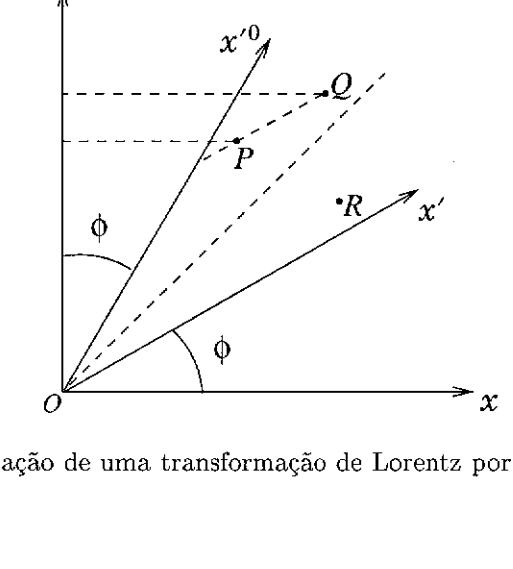


Figura 6.3: Representação de uma transformação de Lorentz por um diagrama de Minkowski.

Considere, agora, um evento O e tome-o como origem de um referencial inercial $\mathcal{K}$. Na Fig. 6.2 o eixo vertical indica $x^0 = ct$ e os outros eixos representam o espaço (a direção z foi suprimida). As retas A e B, inclinadas de 45^o relativamente ao eixo do tempos, representam as *linhas de universo* (trajetórias no espaço-tempo) de raios luminosos viajando em direções opostas, mas cuja emissão constitui o evento O. As linhas de universo dos raios luminosos em toda as direções, cuja emissão corresponde ao evento O, formam um cone com vértice em O, o chamado *cone de luz* de O, que divide o espaço-tempo em quatro regiões. A região I é chamada de *futuro absoluto* de O e um evento E_I de I está separado de O por um intervalo do tipo tempo. A magnitude da projeção horizontal de um vetor como $\overrightarrow{OE_I}$ é a distância espacial entre os eventos, ao passo que a projeção vertical é $c\Delta t$, onde Δt é o intervalo de tempo entre os eventos. Note que $c\Delta t$ é a distância percorrida por um raio luminoso durante o intervalo de tempo entre os eventos. Como a distância espacial entre os eventos O e E_I é menor do que c vezes o intervalo temporal que os separa, pode haver uma relação de causa e efeito entre O e E_I. Por exemplo, o evento O poderia representar uma explosão e o evento E_I representar a chegada do estrondo produzido pela explosão. A região II é o *passado absoluto* de O, e pode haver uma relação de causa e efeito entre E_{II} e O. Por outro lado, é impossível haver conexão causal entre O e um evento nas regiões III ou IV. Por exemplo, a distância percorrida pela luz no intervalo de tempo entre os eventos O e E_{IV} é menor do que a separação espacial entre eles, de modo que para haver relação de causa e efeito entre O e E_{IV} seria necessário existir algum processo físico capaz de propagar-se com velocidade superior à velocidade da luz no vácuo, o que é proibido pela teoria especial da relatividade.

Representações gráficas do espaço-tempo como a da Fig. 6.2, conhecidas como *diagramas de Minkowski*, são úteis para ilustrar algumas das principais propriedades das transformações de Lorentz. Representemos por eixos mutuamente perpendiculares as coordenadas x^0=ct e x do sistema de referência $\mathcal{K}$, como na Fig. 6.3. Consideremos agora o sistema $\mathcal{K}'$ que se move em relação a $\mathcal{K}$ com velocidade v ao longo da direção x. A reta $x' = 0$, que corresponde ao eixo x'^0 de $\mathcal{K}'$, satisfaz a equação $x = \beta x^0$ com $\beta = v/c$, como se deduz de (6.1.1). Portanto, o ângulo ϕ entre os eixos x^0 e x'^0 é menor do que 45^o

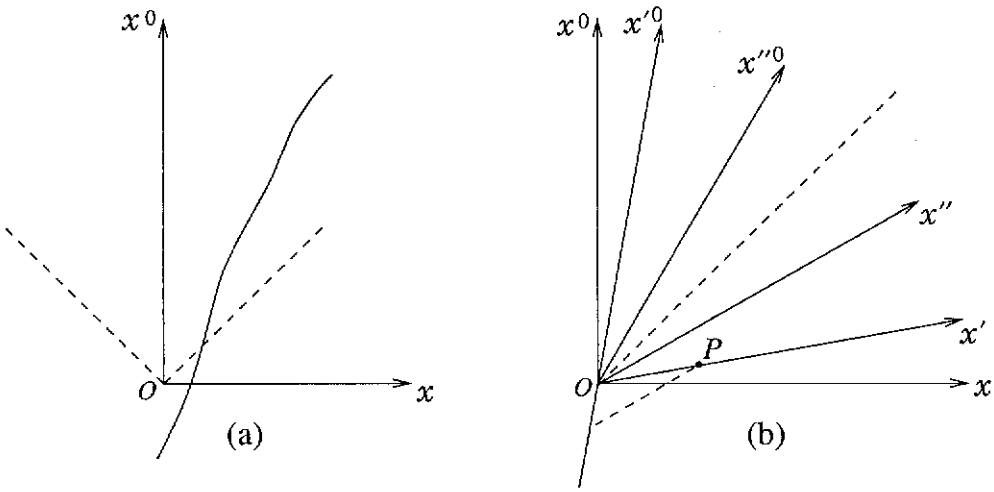


Figura 6.4: Diagramas de Minkowski.

porque tan $\phi = \beta < 1$. Analogamente, o eixo x' é definido por $x'^0 = 0$ e corresponde, pela Eq.(6.1.1), à reta $x^0 = \beta x$, que faz o mesmo ângulo ϕ com o eixo x^0. Uma transformação de Lorentz corresponde a uma transição dos eixos perpendiculares (x^0, x) para os eixos inclinados (x'^0, x'). As coordenadas de um evento são obtidas traçando-se linhas paralelas aos respectivos eixos coordenados. Por intermédio desses diagramas é simples perceber a relatividade da simultaneidade, uma característica marcante da relatividade especial. Dados dois eventos espacialmente separados P e Q que são simultâneos em $\mathcal{K}'$, eles não são simultâneos em $\mathcal{K}$, como a Fig. 6.3 mostra claramente (o segmento $\overline{PQ}$ é paralelo ao eixo x'). Por outro lado, se R é um evento separado de O por um intervalo do tipo espaço, existe um referencial inercial no qual os eventos O e R ocorrem no mesmo instante.

A linha de universo de uma partícula é representada no diagrama de Minkowski por uma curva, como ilustra a Fig. 6.4(a). A curva não é arbitrária, no entanto, porque de $dx/dx^0 = \beta < 1$ conclui-se que a tangente à linha de universo da partícula faz um ângulo com o eixo dos tempos sempre menor do que 45°. Se a trajetória da partícula no espaço-tempo contiver o evento O, sua linha de universo estará inteiramente contida no cone de luz do evento O. A linha de universo de um raio luminoso é uma reta inclinada de 45° em relação aos eixos coordenados. Como exemplo da utilidade dos diagramas do espaço-tempo, mostremos de outra maneira a impossibilidade de conexão causal entre eventos separados por um intervalo do tipo espaço. Na Fig. 6.4(b), suponha que O represente a emissão e P a recepção de um hipotético sinal capaz de propagar-se em $\mathcal{K}$ com velocidade superior à da luz no vácuo. Então existiria um referencial $\mathcal{K}'$ em que a propagação desse sinal seria instantânea, isto é, sua velocidade seria infinita. Pior ainda, haveria um referencial $\mathcal{K}''$ no qual o sinal seria recebido antes de ter sido emitido! Apesar de sua inegável utilidade, faremos uso desses diagramas apenas ocasionalmente porque, em geral, os argumentos algébricos ou analíticos são muito mais poderosos.

6.3 Vetores e Tensores

O princípio da relatividade afirma que as leis da Física têm a mesma forma em todos os referenciais inerciais. A fim de fazer valer este princípio, a expressão matemática das leis físicas deve envolver somente quantidades com regras de transformação bem definidas quando se passa de um referencial inercial para outro.

Definição 6.3.1. Um *escalar* é uma quantidade invariante sob transformações de Lorentz, ou seja, cujo valor é o mesmo em todos os referenciais inerciais.

A carga elétrica de uma partícula, o intervalo entre dois eventos e a velocidade da luz no vácuo são exemplos de escalares.

Definição 6.3.2. Um *quadrivetor* (ou, simplesmente, *vetor*) *contravariante* é um conjunto de quatro quantidades (V^0, V^1, V^2, V^3) que, sob uma transformação de Lorentz, transformam-se da mesma maneira que o deslocamento (dx^0, dx^1, dx^2, dx^3), isto é,

$$V'^{\mu} = \Lambda^{\mu}{}_{\nu} V^{\nu} \ . \tag{6.3.1}$$

Em suma, o protótipo de vetor contravariante é o vetor deslocamento dx^{μ}.

Teorema 6.3.1. Se V^{μ} e W^{μ} são vetores contravariantes, então

$$g_{\mu\nu} V^{\mu} W^{\nu} = V^0 W^0 - V^1 W^1 - V^2 W^2 - V^3 W^3 = \text{escalar} \ . \tag{6.3.2}$$

Este teorema é uma conseqüência imediata da Eq.(6.1.16), e sua validade sugere introduzir uma operação de abaixamento de índice definida por

$$V_{\mu} = g_{\mu\nu} V^{\nu} \ , \tag{6.3.3}$$

pois com esta operação a Eq.(6.3.2) assume a forma usual de um produto escalar.

Definição 6.3.3. O *produto escalar* de dois vetores contravariantes V^{μ} e W^{μ} é definido por[8]

$$\mathbf{V} \cdot \mathbf{W} = V_{\mu} W^{\mu} = V_0 W^0 + V_1 W^1 + V_2 W^2 + V_3 W^3 \ , \tag{6.3.4a}$$

com V_{μ} dado por (6.3.3). Em particular,

$$\mathbf{V}^2 \equiv \mathbf{V} \cdot \mathbf{V} = (V^0)^2 - (V^1)^2 - (V^2)^2 - (V^3)^2 \ . \tag{6.3.4b}$$

■ **Exercício 6.3.1.** Prove que

$$V_{\mu} W^{\mu} = V^{\mu} W_{\mu} \ . \tag{6.3.5}$$

[8]Neste capítulo, letras maiúsculas da segunda metade do alfabeto, em negrito, serão ocasionalmente empregadas para denotar quadrivetores.

Mostre, ainda, que se $V^\mu = (V^0, V^1, V^2, V^3)$ então

$$V_\mu = (V^0, -V^1, -V^2, -V^3) \ , \tag{6.3.6}$$

e o abaixamento de índice apenas troca o sinal das componentes espaciais do quadrivetor. ■

É conveniente introduzir os elementos $g^{\mu\nu}$ da inversa da matriz $\mathbf{G}$ definida por (6.1.14). Mas, como $\mathbf{G}^{-1} = \mathbf{G}$, é imediato que

$$(g^{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{6.3.7}$$

e, evidentemente,

$$g^{\mu\alpha}\, g_{\alpha\nu} = \delta^\mu_\nu \ . \tag{6.3.8}$$

Note, ainda, que $g^{\mu\nu}$ e $g_{\mu\nu}$ são simétricos em seus índices, isto é,

$$g^{\mu\nu} = g^{\nu\mu} \quad , \quad g_{\mu\nu} = g_{\nu\mu} \ . \tag{6.3.9}$$

Com a ajuda de $g^{\mu\nu}$ definimos a operação de levantamento de índice por

$$V^\mu = g^{\mu\nu}\, V_\nu \ . \tag{6.3.10}$$

É claro que o abaixamento do índice de um vetor contravariante V^μ por (6.3.3) seguido de um levantamento do índice por (6.3.10) reproduz o vetor V^μ original.

Se ϕ é uma função escalar, sua diferencial

$$d\phi = \frac{\partial \phi}{\partial x^0}\, dx^0 + \frac{\partial \phi}{\partial x^1}\, dx^1 + \frac{\partial \phi}{\partial x^2}\, dx^2 + \frac{\partial \phi}{\partial x^3}\, dx^3 = \frac{\partial \phi}{\partial x^\mu}\, dx^\mu \tag{6.3.11}$$

também é um escalar. Comparando esta equação com (6.3.4), concluímos que o gradiente de uma função escalar é um tipo de vetor distinto do vetor contravariante, e sua lei de transformação é idêntica à de V_μ definido por (6.3.3) a partir de um vetor contravariante V^μ. Vejamos, portanto, como se transforma V_μ:

$$V'_\mu = g_{\mu\nu}\, V'^\nu = g_{\mu\nu}\, \Lambda^\nu{}_\alpha\, V^\alpha = g_{\mu\nu}\, \Lambda^\nu{}_\alpha\, g^{\alpha\beta}\, V_\beta \equiv \Lambda_\mu{}^\beta\, V_\beta \ , \tag{6.3.12}$$

onde definimos

$$\Lambda_\mu{}^\beta = g_{\mu\nu}\, g^{\beta\alpha}\, \Lambda^\nu{}_\alpha \ . \tag{6.3.13}$$

Definição 6.3.4. Um *vetor covariante* V_μ é definido pela lei de transformação

$$V'_{\mu} = \Lambda_{\mu}{}^{\nu} V_{\nu} \ , \tag{6.3.14}$$

onde os coeficientes $\Lambda_{\mu}{}^{\nu}$ são determinados em termos dos elementos da matriz da transformação de Lorentz por (6.3.13). Em suma, o protótipo de vetor covariante é o gradiente de uma função escalar $\partial_{\mu}\phi \equiv \partial\phi/\partial x^{\mu}$.

■ **Exercício 6.3.3** Fazendo uso das Eqs. (6.1.19) e (6.3.8) mostre que

$$\Lambda_{\mu}{}^{\nu} \Lambda^{\mu}{}_{\alpha} = \delta^{\nu}_{\alpha} \ , \tag{6.3.15}$$

e daí conclua que a matriz $(\Lambda_{\mu}{}^{\nu})$ é a transposta da inversa de $\mathbf{\Lambda} = (\Lambda^{\mu}{}_{\nu})$. ■

Definição 6.3.5. Um *tensor contravariante* de ordem r é um objeto $T^{\alpha_1\alpha_2\cdots\alpha_r}$ com 4^r componentes que, sob uma transformação de Lorentz, transformam-se segundo a regra

$$T'^{\alpha_1\alpha_2\cdots\alpha_r} = \Lambda^{\alpha_1}{}_{\beta_1} \Lambda^{\alpha_2}{}_{\beta_2} \cdots \Lambda^{\alpha_r}{}_{\beta_r} T^{\beta_1\beta_2\cdots\beta_r} \ . \tag{6.3.16a}$$

Analogamente, um *tensor covariante* $Z_{\alpha_1\alpha_2\ldots\alpha_r}$ de ordem r é definido pela lei de transformação

$$Z'_{\alpha_1\alpha_2\ldots\alpha_r} = \Lambda_{\alpha_1}{}^{\beta_1} \Lambda_{\alpha_2}{}^{\beta_2} \ldots \Lambda_{\alpha_r}{}^{\beta_r} Z_{\beta_1\beta_2\cdots\beta_r} \ . \tag{6.3.16b}$$

Tensores mistos, com índices covariantes e contravariantes, definem-se de modo similar.

Exemplos de tensores contravariantes de segunda ordem são

$$V^{\mu}W^{\nu} \quad , \quad V^{\mu}W^{\nu} + V^{\nu}W^{\mu} \quad , \quad V^{\mu}W^{\nu} - V^{\nu}W^{\mu} \ , \tag{6.3.17}$$

onde V^{μ} e W^{μ} são vetores contravariantes.

As operações de abaixamento e levantamento de índices são aplicáveis a um ou mais índices de tensores de ordem arbitrária:

$$T_{\mu}{}^{\nu} = g_{\mu\alpha} T^{\alpha\nu} \quad , \quad T^{\mu}{}_{\nu} = g_{\nu\alpha} T^{\mu\alpha} \ , \tag{6.3.18a}$$

$$T_{\mu\nu} = g_{\mu\alpha} g_{\nu\beta} T^{\alpha\beta} \quad , \quad T^{\mu\nu} = g^{\mu\alpha} g^{\nu\beta} T_{\alpha\beta} \ . \tag{6.3.18b}$$

O produto escalar de um vetor por um tensor reduz de uma unidade a ordem do tensor. Por exemplo:

$$V_{\mu} T^{\alpha\beta\mu} = \text{tensor de segunda ordem} \ ; \tag{6.3.19a}$$

$$V_{\mu} Y^{\mu\nu} = \text{tensor de primeira ordem} = \text{vetor} \ . \tag{6.3.19b}$$

Verifiquemos, por exemplo, esta última afirmação. Definindo $W^\nu = V_\mu Y^{\mu\nu}$ e usando as Eqs.(6.3.14) e (6.3.15), temos

$$W'^\nu = V'_\mu Y'^{\mu\nu} = \Lambda_\mu{}^\alpha V_\alpha \Lambda^\mu{}_\beta \Lambda^\nu{}_\gamma Y^{\beta\gamma} = \Lambda_\mu{}^\alpha \Lambda^\mu{}_\beta V_\alpha \Lambda^\nu{}_\gamma Y^{\beta\gamma}$$
$$= \delta^\alpha_\beta V_\alpha \Lambda^\nu{}_\gamma Y^{\beta\gamma} = V_\beta \Lambda^\nu{}_\gamma Y^{\beta\gamma} = \Lambda^\nu{}_\gamma (V_\beta Y^{\beta\gamma}) = \Lambda^\nu{}_\gamma W^\gamma \ , \tag{6.3.20}$$

o que comprova o caráter vetorial de W^ν.

Uma operação do tipo

$$T^\mu{}_\mu{}^\alpha \quad , \quad T^{\mu\nu}{}_\nu \quad , \quad T^{\mu\nu}{}_\mu \ , \tag{6.3.21}$$

é chamada de *contração*. Assim, por exemplo, $T^\mu{}_\mu{}^\alpha$ é a contração dos dois primeiros índices de $T^{\mu\nu\alpha}$. Claramente, a contração reduz a ordem do tensor em duas unidades. Por exemplo, no caso de um tensor de segunda ordem $T^{\mu\nu}$, o escalar[9]

$$T^\mu{}_\mu = T_\mu{}^\mu \tag{6.3.22}$$

é chamado de traço do tensor.

Notemos, finalmente, que a classificação de intervalos aplica-se a vetores:

$$\begin{aligned} V_\mu V^\mu > 0 \quad &\Longleftrightarrow \quad V^\mu \text{ é vetor do tipo tempo} \ ; \\ V_\mu V^\mu < 0 \quad &\Longleftrightarrow \quad V^\mu \text{ é vetor do tipo espaço} \ ; \\ V_\mu V^\mu = 0 \quad &\Longleftrightarrow \quad V^\mu \text{ é vetor do tipo luz} \ . \end{aligned} \tag{6.3.23}$$

Esta classificação é absoluta porque $V_\mu V^\mu$ é invariante sob transformações de Lorentz.

6.4 Campos Tensoriais

Definição 6.4.1. Um campo escalar é uma função $\phi(P)$ cujo valor em cada ponto P do espaço-tempo é invariante sob transformações de Lorentz, isto é, $\phi'(P) = \phi(P)$. Se o ponto P tem coordenadas $x \equiv (x^0, x^1, x^2, x^3)$ no sistema de referência inercial $\mathcal{K}$ e $x' \equiv (x'^0, x'^1, x'^2, x'^3)$ no sistema de referência inercial $\mathcal{K}'$, temos

$$\phi'(x') = \phi(x) \ . \tag{6.4.1}$$

Definição 6.4.2. Um campo tensorial $T^{\mu\nu\cdots\rho}(P)$ é um tensor associado a cada ponto P do espaço-tempo e caracterizado pela lei de transformação

$$T'^{\mu\nu\cdots\rho}(x') = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \cdots \Lambda^\rho{}_\sigma T^{\alpha\beta\cdots\sigma}(x) \ . \tag{6.4.2}$$

O operador nabla quadridimensional

[9] Um escalar pode ser considerado um tensor de ordem zero.

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = (\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}) = (\frac{1}{c}\frac{\partial}{\partial t}, \nabla) \tag{6.4.3}$$

é um vetor covariante. Isto deduz-se do fato de $\partial_\mu \phi$ ser um vetor covariante, como vimos na Seção 6.3, e da definição de campo escalar. De fato, combinando

$$(\partial_\mu \phi)' = \Lambda_\mu{}^\nu (\partial_\nu \phi) \quad \text{ou} \quad \partial'_\mu \phi' = \Lambda_\mu{}^\nu \partial_\nu \phi \tag{6.4.4}$$

com $\phi' = \phi$, concluímos que, como operador,

$$\partial'_\mu = \Lambda_\mu{}^\nu \partial_\nu \ , \tag{6.4.5}$$

confirmando a asserção inicial. Analogamente,

$$\partial^\mu = g^{\mu\nu}\partial_\nu = \frac{\partial}{\partial x_\mu} = (\frac{\partial}{\partial x_0}, \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}) = (\frac{1}{c}\frac{\partial}{\partial t}, -\nabla) \tag{6.4.6}$$

é um vetor contravariante.

Campos tensoriais podem ter sua ordem aumentada por diferenciação. Por exemplo, se ϕ é um campo escalar então

$$\partial_\mu \partial_\nu \phi = \frac{\partial^2 \phi}{\partial x^\mu \partial x^\nu} = \text{campo tensorial covariante de segunda ordem} \ . \tag{6.4.7}$$

Se $T^{\mu\nu}$ é um campo tensorial contravariante de segunda ordem segue-se que

$$\partial^\rho T^{\mu\nu} = \text{campo tensorial contravariante de terceira ordem} \ . \tag{6.4.8}$$

Havendo contração de índices, por diferenciação pode-se reduzir a ordem de um campo tensorial. Por exemplo, se $F^{\mu\nu}$ é um campo tensorial contravariante de segunda ordem, verifica-se imediatamente que

$$\partial_\nu F^{\mu\nu} = \text{campo vetorial contravariante} \ . \tag{6.4.9}$$

Divergência, Rotacional e d'Alembertiano

Definição 6.4.3. A *divergência* de um campo vetorial V^μ é um campo escalar denotado por $\operatorname{div} V$ e definido por

$$\operatorname{div} V = \partial_\mu V^\mu \ . \tag{6.4.10}$$

No caso particular de espaço tridimensional, o rotacional de um campo vetorial é também um campo vetorial. Em todos os demais casos, o rotacional de um campo vetorial é naturalmente definido como um campo tensorial anti-simétrico de segunda ordem. A excepcionalidade do espaço tridimensional é fácil de entender porque apenas em três

dimensões há uma correspondência biunívoca entre vetores e matrizes anti-simétricas, como vimos na Seção 3.4.

Definição 6.4.4. O *rotacional* de um campo vetorial V^μ é um campo tensorial de segunda ordem denotado por $(\mathrm{rot}\, V)^{\mu\nu}$ e definido por

$$(\mathrm{rot}\, V)^{\mu\nu} = \partial^\mu V^\nu - \partial^\nu V^\mu \ . \tag{6.4.11}$$

A divergência do gradiente constitui o operador de d'Alembert ou "d'Alembertiano", que é a versão quadridimensional do laplaciano:

$$\Box \equiv \partial_\mu \partial^\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 \ . \tag{6.4.12}$$

Este operador é um escalar, de modo que sua aplicação não altera a ordem de um campo tensorial.

6.5 Leis Físicas em Forma Covariante

O aparato tensorial desenvolvido nas seções anteriores tem como objetivo permitir a expressão matemática das leis da Física em forma *covariante*, isto é, *válida em todos os referenciais inerciais*. Suponha que, num dado sistema de referência inercial $\mathcal{K}$, uma lei da Física possa ser expressa na forma

$$T^{\mu\nu\cdots\rho} = 0 \ , \tag{6.5.1}$$

onde $T^{\mu\nu\cdots\rho}$ é um tensor ou campo tensorial. Em virtude das leis de transformação homogêneas (6.3.16) ou (6.4.2) de suas componentes, se um tensor ou campo tensorial é nulo num sistema inercial ele será nulo em todos os demais sistemas inerciais. Em outras palavras, num sistema inercial $\mathcal{K}'$ teremos

$$T'^{\mu\nu\cdots\rho} = 0 \ , \tag{6.5.2}$$

se (6.5.1) valer em $\mathcal{K}$. Em conseqüência, a lei física descrita por (6.5.1) em $\mathcal{K}$ tem a mesma forma em *todos* os referenciais inerciais. Dizemos que uma lei física representada por uma equação do tipo (6.5.1) está expressa numa forma *manifestamente covariante*, bastando olhar para a equação para se reconhecer sua validade em todos os sistemas de referência inerciais.

Dado um conjunto de leis físicas sabidamente válidas em todos os sistemas inerciais, é importante exprimi-las numa forma manifestamente covariante.[10] Isso, entre outras vantagens, facilita a determinação de como se transformam as grandezas físicas relevantes quando se passa de um referencial inercial para outro. Um exemplo de suma importância é fornecido pela eletrodinâmica clássica.

[10] Quando se refere a uma equação ou lei física, a palavra "covariante" significa "válida com a mesma forma em todos os referenciais inerciais".

Formulação Covariante da Eletrodinâmica

Em unidades *CGS* gaussianas as equações de Maxwell no vácuo escrevem-se

$$\nabla \cdot \mathbf{E} = 4\pi\rho \quad , \quad \nabla \cdot \mathbf{B} = 0 \quad , \quad \nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = 0 \quad , \quad \nabla \times \mathbf{B} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c}\mathbf{j} \ . \tag{6.5.3}$$

Estas equações prevêem que a luz se propaga no vácuo com velocidade constante c. Segundo o postulado fundamental da teoria especial da relatividade, isto é verdadeiro em todos os sistemas de referência inerciais, o que é compatível com as transformações de Lorentz mas contradiz as transformações de Galileu. Portanto, as Eqs.(6.5.3) devem ser válidas em todos os referenciais inerciais, devendo ser possível reescrevê-las em forma manifestamente covariante *sem alterar seu conteúdo.*

Comecemos pela lei da conservação local da carga elétrica

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0 \ , \tag{6.5.4}$$

que se deduz facilmente das Eqs.(6.5.3). Vamos introduzir o *candidato* a quadrivetor

$$J^\mu = (c\rho \, , \, j_x \, , \, j_y \, , \, j_z) \equiv (c\rho \, , \, \mathbf{j}) \ , \tag{6.5.5}$$

que permite reescrever (6.5.4) na forma

$$\partial_\mu J^\mu = 0 \ . \tag{6.5.6}$$

Como a experiência indica a validade da lei da conservação da carga elétrica em todos os sistemas inerciais, conclui-se que J^μ é um quadrivetor. Pela Definição 6.4.3, a Eq.(6.5.6) assevera que um escalar é igual a zero, o que é verdadeiro em todos os referenciais inerciais. Assim, (6.5.6) é a forma manifestamente covariante da equação da continuidade (6.5.4). O quadrivetor J^μ definido por (6.5.5) é chamado de *quadricorrente* ou *quadrivetor densidade de carga-corrente.*

■ **Exercício 6.5.1.** Prove que as densidades de carga e corrente transformam-se de acordo com

$$\rho' = \gamma\left(\rho - \frac{v}{c^2} j_x\right) \quad , \quad j'_x = \gamma\left(j_x - v\,\rho\right) \quad , \quad j'_y = j_y \quad , \quad j'_z = j_z \ , \tag{6.5.7}$$

no caso da transformação de Lorentz pura (6.1.1). ■

É útil introduzir os potenciais ϕ e $\mathbf{A}$ tais que

$$\mathbf{B} = \nabla \times \mathbf{A} \quad , \quad \mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \ . \tag{6.5.8}$$

Consideremos, então,

$$A^\mu = (\phi \, , \, \mathbf{A}) \tag{6.5.9}$$

como novo candidato a quadrivetor. A primeira das equações (6.5.8) sugere introduzir um candidato a tensor de segunda ordem $F^{\mu\nu}$ definido como o rotacional de A^μ:

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \ . \tag{6.5.10}$$

É fácil comprovar que as componentes de $F^{\mu\nu}$ são

$$(F^{\mu\nu}) = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \ . \tag{6.5.11}$$

A título de verificação consideremos, por exemplo, F^{01}:

$$F^{01} = \partial^0 A^1 - \partial^1 A^0 = \partial_0 A_x + \partial_1 \phi = \frac{1}{c}\frac{\partial A_x}{\partial t} + \frac{\partial \phi}{\partial x} = -E_x \ .$$

■ **Exercício 6.5.2.** Verifique que as demais componentes de $F^{\mu\nu}$ são dadas corretamente pela Eq.(6.5.11). ■

Uma vez que suas componentes exprimem-se em termos das componentes de $\mathbf{E}$ e $\mathbf{B}$, $F^{\mu\nu}$ é chamado de *tensor do campo eletromagnético*. É importante destacar que $F^{\mu\nu}$ é um tensor anti-simétrico:

$$F^{\mu\nu} = -F^{\nu\mu} \ . \tag{6.5.12}$$

A primeira e a quarta equações de Maxwell (6.5.3) sugerem uma igualdade tendo à esquerda a divergência de $F^{\mu\nu}$ e, à direita, J^μ. É necessária uma análise mais cuidadosa para se descobrir a forma covariante adequada das outras duas equações. O resultado é que as equações de Maxwell em forma covariante são

$$\partial_\mu F^{\mu\nu} = \frac{4\pi}{c} J^\nu \ , \tag{6.5.13}$$

$$\partial^\rho F^{\mu\nu} + \partial^\mu F^{\nu\rho} + \partial^\nu F^{\rho\mu} = 0 \ . \tag{6.5.14}$$

As equações (6.5.13) coincidem com as equações de Maxwell não-homogêneas, ao passo que as equações (6.5.14) – apenas quatro delas são independentes – são idênticas às equações de Maxwell homogêneas. Por exemplo, tome $\nu = 0$ em (6.5.13):

$$\partial_\mu F^{\mu 0} = \frac{4\pi}{c} J^0 \implies \frac{\partial F^{00}}{\partial x^0} + \frac{\partial F^{10}}{\partial x^1} + \frac{\partial F^{20}}{\partial x^2} + \frac{\partial F^{30}}{\partial x^3} = \frac{4\pi}{c} c\rho \implies \boldsymbol{\nabla}\cdot\mathbf{E} = 4\pi\rho \ .$$

Analogamente, escolhendo $\rho = 0$, $\mu = 1$, $\nu = 2$ em (6.5.14) resulta:

$$0 = \partial^0 F^{12} + \partial^1 F^{20} + \partial^2 F^{01} = \frac{1}{c}\frac{\partial(-B_z)}{\partial t} - \frac{\partial E_y}{\partial x} - \frac{\partial(-E_x)}{\partial y} \implies \left(\boldsymbol{\nabla}\times\mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}\right)_z = 0 \ .$$

A validade empiricamente corroborada das equações de Maxwell em todos os referenciais inerciais nos leva a concluir que efetivamente A^μ é um quadrivetor e $F^{\mu\nu}$ um tensor de segunda ordem. Assim, (6.5.13) e (6.5.14) constituem a expressão das equações de Maxwell em forma manifestamente covariante.

■ **Exercício 6.5.3.** Usando a lei de transformação (6.4.2) de um tensor contravariante, mostre que, no caso da transformação de Lorentz pura (6.1.12), o campo eletromagnético se transforma conforme

$$E'_x = E_x \quad , \qquad B'_x = B_x \ , \tag{6.5.15a}$$

$$E'_y = \gamma\,(E_y - \beta B_z) \quad , \qquad B'_y = \gamma\,(B_y + \beta E_z) \ , \tag{6.5.15b}$$

$$E'_z = \gamma\,(E_z + \beta B_y) \quad , \qquad B'_z = \gamma\,(B_z - \beta E_y) \ . \tag{6.5.15c}$$

Sugestão: denotando por $\mathcal{F}$ a matriz (6.5.11), mostre que a lei de transformação de $F^{\mu\nu}$ pode ser escrita na forma matricial $\mathcal{F}' = \Lambda\mathcal{F}\Lambda^T$. ■

6.6 Dinâmica Relativística

A mecânica newtoniana é invariante sob transformações de Galileu. Em particular, a segunda lei de Newton para uma partícula,

$$m\frac{d^2\mathbf{r}}{dt^2} = \mathbf{F} \ , \tag{6.6.1}$$

preserva sua forma admitindo-se a invariância de $\mathbf{F}$, já que a aceleração é invariante sob transformações galileanas. Para dar à segunda lei de Newton uma forma invariante sob transformações de Lorentz, é de se esperar que seja necessário modificar o seu conteúdo, ao contrário do que ocorreu com a teoria eletromagnética, cuja covariância relativística já estava embutida nas equações tal como originalmente propostas por Maxwell.

Quadrivelocidade, Quadriaceleração e Quadrimomento

A fim de obter a versão quadridimensional da Eq.(6.6.1), comecemos procurando definir a quadrivelocidade. A idéia mais imediata seria considerar dx^μ/dt, mas este objeto não é um quadrivetor porque dt não é um escalar. Devemos, portanto, substituir dt por um intervalo de tempo que seja um escalar. O *intervalo de tempo próprio* $d\tau$ é definido como o intervalo de tempo medido no sistema de referência em que a partícula encontra-se instantaneamente em repouso. Suponha que, no instante t de um referencial $\mathcal{K}$, a partícula tenha velocidade $\mathbf{v}$ quando vista de $\mathcal{K}$. Seja $\mathcal{K}'$ um referencial inercial que move-se com velocidade $\mathbf{v}$ relativamente a $\mathcal{K}$, de modo que, do ponto de vista de $\mathcal{K}'$, a partícula encontra-se em repouso no instante t de $\mathcal{K}$. Seja ds o intervalo invariante entre os eventos definidos pela passagem da partícula pelos pontos (x, y, z) no instante t e $(x + dx, y + dy, z + dz)$ no instante $t + dt$. Evidentemente, $d\mathbf{r} = \mathbf{v}dt$, de sorte que, no referencial $\mathcal{K}$,

$$ds^2 = c^2\,dt^2 - d\mathbf{r}\cdot d\mathbf{r} = c^2\,dt^2\,(1 - v^2/c^2) \ . \tag{6.6.2}$$

Mas, do ponto de vista de $\mathcal{K}'$,

$$ds'^2 = c^2\, d\tau^2 \tag{6.6.3}$$

porque no intervalo de tempo $d\tau$ medido em $\mathcal{K}'$ a partícula permaneceu imóvel, visto que o referencial $\mathcal{K}'$ acompanhou o movimento da partícula. Mas $ds^2 = ds'^2$ por causa da invariância do intervalo, donde se conclui que

$$d\tau = \sqrt{1 - v^2/c^2}\, dt \equiv dt/\gamma \ , \tag{6.6.4}$$

com γ definido por (6.1.4). Note que a Eq.(6.6.3) comprova que $d\tau$ é um escalar, pois ds' e a velocidade da luz no vácuo são escalares.

A quadrivelocidade U^μ é o quadrivetor definido por

$$U^\mu = \frac{dx^\mu}{d\tau} \ , \tag{6.6.5}$$

cujas componentes são

$$U^\mu = (\gamma\, c\, ,\, \gamma\, \mathbf{v}) \ , \tag{6.6.6}$$

onde $\mathbf{v} = d\mathbf{r}/dt$ é a velocidade tridimensional da partícula. Uma conseqüência imediata desta definição é que o quadrado de U^μ é sempre uma mesma constante independentemente do estado de movimento da partícula:

$$U^\mu U_\mu = c^2 \ . \tag{6.6.7}$$

Isto mostra que a quadrivelocidade U^μ é um vetor do tipo tempo que nunca se anula.

Analogamente, a quadriaceleração é o quadrivetor definido por

$$\mathcal{A}^\mu = \frac{dU^\mu}{d\tau} = \frac{d^2x^\mu}{d\tau^2} \ , \tag{6.6.8}$$

cujas componentes são

$$\mathcal{A}^\mu = (\gamma^4 \frac{\mathbf{v}\cdot\mathbf{a}}{c} \ ,\ \gamma^4 \frac{(\mathbf{v}\cdot\mathbf{a})}{c^2}\mathbf{v} + \gamma^2\mathbf{a}) \ , \tag{6.6.9}$$

onde $\mathbf{a} = d\mathbf{v}/dt$ é a aceleração tridimensional. No referencial de repouso momentâneo da partícula, $\mathbf{v} = 0$ e as componentes da quadriaceleração reduzem-se a $\mathcal{A}^\mu = (0, \mathbf{a})$, donde se deduz que a quadriaceleração é um vetor do tipo espaço e só se anula se a *aceleração própria* – isto é, a aceleração medida no referencial de repouso instantâneo – for zero.

■ **Exercício 6.6.1.** Prove diretamente, ou utilizando (6.6.7), que

$$\mathcal{A}^\mu\, U_\mu = 0 \ , \tag{6.6.10}$$

isto é, a quadrivelocidade e a quadriaceleração são vetores ortogonais. ■

Finalmente, o quadrimomento de uma partícula define-se por

$$P^\mu = mU^\mu = (\gamma\, m\, c\,,\, \gamma\, m\, \mathbf{v}) \ , \tag{6.6.11}$$

onde m é um escalar chamado de *massa de repouso* ou, simplesmente, massa da partícula. De (6.6.7) decorre imediatamente

$$P^\mu P_\mu = m^2 c^2 \ . \tag{6.6.12}$$

Podemos escrever

$$P^\mu = (\gamma\, m\, c\,,\, \mathbf{p}) \ , \tag{6.6.13}$$

onde

$$\mathbf{p} = \gamma\, m\, \mathbf{v} = \frac{m\,\mathbf{v}}{\sqrt{1 - v^2/c^2}} \tag{6.6.14}$$

é o *momento linear relativístico*. O momento linear relativístico reduz-se à expressão newtoniana $m\mathbf{v}$ no limite $v \ll c$ e uma experiência imaginária, concebida por G. N. Lewis e R. C. Tolman, prova que é exatamente a forma (6.6.14) que garante a validade da lei da conservação do momento linear em todos os referenciais inerciais (Pauli 1958; Bergmann 1976).

Equação de Movimento Covariante

Estamos prontos, agora, para propor a versão covariante da segunda lei de Newton na forma

$$\frac{dP^\mu}{d\tau} = \mathcal{F}^\mu \ , \tag{6.6.15}$$

onde $\mathcal{F}^\mu$ é a quadriforça, também chamada de força de Minkowski. Por razões de consistência, o quadrivetor força $\mathcal{F}^\mu$ tem que ser ortogonal a P^μ:

$$\mathcal{F}^\mu P_\mu = 0 \ . \tag{6.6.16}$$

■ **Exercício 6.6.2.** Demonstre (6.6.16) a partir de (6.6.12). ■

A Eq.(6.6.15) é desprovida de conteúdo a menos que se saiba como relacionar $\mathcal{F}^\mu$ com a força tridimensional $\mathbf{F}$. O procedimento mais simples possível consiste em fazer a hipótese de que permanece válida a conexão newtoniana entre a força e a taxa de variação temporal do momento linear:

$$\frac{d\mathbf{p}}{dt} \equiv \frac{d}{dt}\Big(\frac{m\,\mathbf{v}}{\sqrt{1-v^2/c^2}}\Big) = \mathbf{F} \ . \tag{6.6.17}$$

Comparando esta equação com as componentes espaciais de (6.6.15) e utilizando (6.6.4), encontra-se $\mathcal{F}^\mu = (\mathcal{F}^0, \gamma\mathbf{F})$. A componente temporal $\mathcal{F}^0$ obtém-se facilmente de (6.6.16):

$$\gamma mc\mathcal{F}^0 - \gamma^2 m\mathbf{F}\cdot\mathbf{v} = 0 \quad \Longrightarrow \quad \mathcal{F}^0 = \frac{\gamma}{c}\mathbf{F}\cdot\mathbf{v} \ . \tag{6.6.18}$$

Assim, a força de Minkowski $\mathcal{F}^\mu$ é dada por

$$\mathcal{F}^\mu = \big(\frac{\gamma}{c}\mathbf{F}\cdot\mathbf{v} \ , \ \gamma\mathbf{F}\big) \ , \tag{6.6.19}$$

o que demonstra que $\mathcal{F}^\mu$ tem apenas três componentes independentes.

A consistência de (6.6.19) pode ser testada no caso eletromagnético. A força de Lorentz

$$\mathbf{F} = e\Big(\mathbf{E} + \frac{\mathbf{v}}{c}\times\mathbf{B}\Big) \tag{6.6.20}$$

depende linearmente do campo eletromagnético e da velocidade da partícula. Assim, a força de Minkowski $\mathcal{F}^\mu$ deve ser um quadrivetor construído exclusivamente com $F^{\mu\nu}$ e U^μ, devendo ser linear nestas quantidades. A única forma possível para $\mathcal{F}^\mu$ é

$$\mathcal{F}^\mu = \kappa\, F^{\mu\nu}\, U_\nu \quad , \quad \kappa = \text{escalar} \ , \tag{6.6.21}$$

e a equação de movimento covariante (6.6.15) toma a forma

$$m\mathcal{A}^\mu = \kappa\, F^{\mu\nu}\, U_\nu \ . \tag{6.6.22}$$

No referencial de repouso instantâneo $\mathcal{K}'$ a velocidade da partícula é zero, não há correções relativísticas e a equação de movimento newtoniana (6.6.1) é exata. Mas em $\mathcal{K}'$ temos $\mathcal{A}^\mu = (0, \mathbf{a}')$, $U_\mu = (c, 0, 0, 0)$, o que reduz a componente temporal de (6.6.22) à identidade $0 = 0$, ao passo que parte espacial, obtida fazendo sucessivamente $\mu = 1, 2, 3$, coincide com

$$m\mathbf{a}' = \kappa\, c\, \mathbf{E}' \ , \tag{6.6.23}$$

tendo sido utilizada a Eq.(6.5.11). Como a força de Lorentz no referencial de repouso instantâneo é $\mathbf{F}' = e\mathbf{E}'$, o escalar κ fica determinado: $\kappa = e/c$. Portanto, a equação de movimento covariante de uma partícula carregada num campo eletromagnético externo é

$$\frac{dP^\mu}{d\tau} = \frac{e}{c}\, F^{\mu\nu}\, U_\nu \ . \tag{6.6.24}$$

■ **Exercício 6.6.3.** Mostre que as componentes da força de Minkowski eletromagnética

$$\mathcal{F}^{\mu} = \frac{e}{c} F^{\mu\nu} U_{\nu} \tag{6.6.25}$$

coincidem com (6.6.19), onde $\mathbf{F}$ é a força de Lorentz (6.6.20). ■

Energia Relativística

Por construção, a parte espacial da equação de movimento covariante (6.6.15) coincide com a Eq.(6.6.17), que é o que se chama comumente de equação de movimento relativística da partícula. A pergunta que naturalmente se coloca é: qual é o significado físico da componente temporal da equação de movimento covariante? Com a ajuda de (6.6.4) e (6.6.19) podemos escrever

$$\frac{dP^0}{d\tau} \equiv \gamma \frac{d}{dt}(mc\gamma) = \mathcal{F}^0 = \frac{\gamma}{c}\mathbf{F}\cdot\mathbf{v} \ , \tag{6.6.26}$$

ou seja,

$$\frac{d}{dt}\left(\frac{m\,c^2}{\sqrt{1-v^2/c^2}}\right) = \mathbf{F}\cdot\mathbf{v} \ . \tag{6.6.27}$$

Tendo em conta que $\mathbf{F}\cdot\mathbf{v}$ é a potência fornecida à partícula pela força $\mathbf{F}$, devemos tomar

$$E = \frac{m\,c^2}{\sqrt{1-v^2/c^2}} \tag{6.6.28}$$

como definição de energia relativística da partícula: a componente temporal da equação de movimento covariante representa a lei da conservação da energia.

Em face da definição (6.6.28) podemos escrever

$$P^{\mu} = \left(\frac{E}{c}\,,\,\mathbf{p}\right) \ . \tag{6.6.29}$$

Desta forma, energia e momento linear passam a constituir um quadrivetor, fazendo com que as leis da conservação da energia e do momento linear deixem de ser independentes, tornando-se aspectos de uma lei de conservação covariante do quadrimomento.

No limite não-relativístico, a expansão binomial aplicada a (6.6.28) fornece

$$E = mc^2 + \frac{1}{2}mv^2 + \frac{3}{8}m\frac{v^4}{c^2} + \cdots \ . \tag{6.6.30}$$

O segundo termo é a energia cinética newtoniana, e os termos seguintes são correções relativísticas à expressão clássica. O primeiro termo não tem análogo clássico e implica que uma partícula em repouso contém energia dada pela célebre equação de Einstein

$$E = mc^2 \ , \tag{6.6.31}$$

que estabelece a equivalência entre massa e energia: uma massa m em repouso pode ser convertida numa quantidade equivalente de energia mc^2.

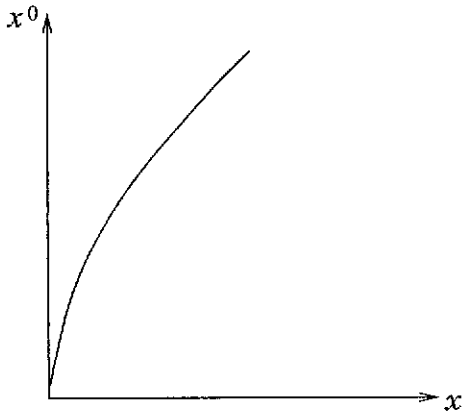


Figura 6.5: Linha de universo de uma partícula relativística sob a ação de uma força constante ("movimento hiperbólico").

■ **Exemplo 6.6.1.** Estudar o movimento unidimensional de uma partícula relativística sujeita a uma força constante, também conhecido como "movimento hiperbólico".

Solução. Este exemplo pode corresponder à aceleração de uma partícula carregada por um campo elétrico constante e uniforme. No caso unidimensional, a Eq.(6.6.17) assume a forma

$$\frac{d}{dt}\left(\frac{m\,\dot{x}}{\sqrt{1-\dot{x}^2/c^2}}\right) = F \ . \tag{6.6.32}$$

Uma integração imediata fornece

$$\frac{\dot{x}}{\sqrt{1-\dot{x}^2/c^2}} = at + b \quad , \quad a = \frac{F}{m} \ , \tag{6.6.33}$$

onde b é uma constante de integração arbitrária. Resolvendo esta equação para $\dot{x}$ somos conduzidos a

$$\dot{x} = \frac{at+b}{\sqrt{1+(at+b)^2/c^2}} \ , \tag{6.6.34}$$

que mostra que, diferentemente do problema não-relativístico, a aceleração da partícula não é constante.[11] Após uma nova integração elementar encontramos

$$x - d = \frac{c^2}{a}\sqrt{1+\left(\frac{at+b}{c}\right)^2} \ . \tag{6.6.35}$$

Consideremos o caso particularmente simples em que a partícula começa em repouso na origem, isto é, $x(0) = 0$ e $\dot{x}(0) = 0$. Em tais circunstâncias $b = 0$ e $d = -c^2/a$, donde

$$x = \frac{c^2}{a}\left[-1+\sqrt{1+\left(\frac{at}{c}\right)^2}\right] \ . \tag{6.6.36}$$

Esta última equação pode ser expressa na forma mais elegante

[11] A *aceleração própria* da partícula é constante e igual a F/m (Problema 6.13).

$$\left(x + \frac{c^2}{a}\right)^2 - c^2t^2 = \frac{c^4}{a^2} \ , \tag{6.6.37}$$

que é a equação de uma hipérbole no plano x^0x (vide Fig. 6.5). ■

■ **Exercício 6.6.4.** Inspecionando a Eq.(6.6.34), prove que $\dot{x} < c$ para qualquer t e que $\dot{x} \to c$ no limite assintótico $t \to \infty$. Encontre o limite não-relativístico da Eq.(6.6.36) passando formalmente ao limite $c \to \infty$ e mostre que os resultados esperados são obtidos. ■

Partículas de Massa Zero

Um importante resultado é o que relaciona a energia da partícula com o seu momento linear, a saber:

$$E = \sqrt{p^2c^2 + m^2c^4} \quad , \quad p^2 = \mathbf{p} \cdot \mathbf{p} \ . \tag{6.6.38}$$

Uma outra característica da mecânica relativística não partilhada pela mecânica newtoniana é a possibilidade de existência de partículas de massa de repouso zero (fótons, grávitons). Fazendo $m = 0$ em (6.6.38) resulta

$$E = |\mathbf{p}|c \ , \tag{6.6.39}$$

que é a conexão relativística entre E e $\mathbf{p}$ para partículas de massa zero. Uma faceta fundamental das partículas de massa nula é que, em módulo, sua velocidade é sempre c em qualquer referencial inercial, como o leitor é instado a demonstrar.

■ **Exercício 6.6.5.** A partir das Eqs.(6.6.12) e (6.6.29) deduza (6.6.38). Prove, ainda, que a velocidade da partícula pode ser expressa na forma

$$\mathbf{v} = \frac{c^2\mathbf{p}}{E} \ , \tag{6.6.40}$$

que não depende da massa. Conclua que $|\mathbf{v}| = c$ se a partícula tiver massa zero. ■

6.7 Colisões Relativísticas

As leis de conservação da energia e do momento linear numa colisão de partículas podem ser sintetizadas na equação covariante

$$\mathbf{P}^{(I)}_{(1)} + \cdots \mathbf{P}^{(I)}_{(N)} = \mathbf{P}^{(F)}_{(1)} + \cdots + \mathbf{P}^{(F)}_{(M)} \ . \tag{6.7.1}$$

O estado inicial (antes da colisão) contém N partículas e o estado final (depois da colisão) contém M partículas, e $\mathbf{P}^{(I)}_{(k)}$ e $\mathbf{P}^{(F)}_{(k)}$ denotam os quadrimomentos inicial e final, respectivamente, da k-ésima partícula. Sendo uma equação tensorial, (6.7.1) assegura a validade

da conservação da energia e do momento linear em todos os sistemas de referência inerciais. Ilustraremos, a seguir, a notável eficiência da Eq.(6.7.1), aliada à invariância do produto escalar (6.3.4), para extrair de forma elegante e concisa informações importantes a respeito de processos envolvendo colisões de partículas com altas energias.

■ **Exemplo 6.7.1.** Antiprótons podem ser produzidos numa colisão próton-próton segundo a reação

$$p + p \rightarrow p + p + p + \bar{p} \ , \tag{6.7.2}$$

onde p denota um próton e $\bar{p}$ um antipróton. Se o próton-alvo encontra-se em repouso no sistema de referência do laboratório, qual é a energia mínima (liminar) do próton incidente para que a reação seja possível?

Solução. De acordo com (6.6.28), uma partícula livre tem a menor energia possível quando está em repouso. Portanto, a energia mínima do próton incidente é aquela que permite a formação dos produtos da reação em repouso. No entanto, os produtos da reação não podem ser formados em repouso no referencial do laboratório, pois isso violaria a lei da conservação do momento linear. Em conseqüência, os produtos da reação só podem ser formados em repouso no referencial do centro de massa, pois é nesse referencial que o momento linear total é zero. Denotando por $\mathbf{P}^{(F)}$ o quadrimomento total dos produtos da colisão, suas componentes no referencial do centro de massa são

$$\mathbf{P}^{(F)} = (4mc, 0, 0, 0) \qquad \text{(no referencial do centro de massa)} \ , \tag{6.7.3}$$

porque cada um dos produtos da reação tem quadrimomento com componentes $(mc, 0, 0, 0)$, onde m é a massa do próton (igual à do antipróton). A lei de conservação do quadrimomento escreve-se

$$\mathbf{P} + \mathbf{Q} = \mathbf{P}^{(F)} \ , \tag{6.7.4}$$

onde $\mathbf{P}$ e $\mathbf{Q}$ são, respectivamente, o quadrimomento do próton incidente e o quadrimomento do próton-alvo. Quadrando a Eq.(6.7.4) resulta

$$\mathbf{P}^2 + 2\mathbf{P}\cdot\mathbf{Q} + \mathbf{Q}^2 = \mathbf{P}^{(F)^2} \quad \Longrightarrow \quad m^2c^2 + 2\mathbf{P}\cdot\mathbf{Q} + m^2c^2 = \mathbf{P}^{(F)^2} \ , \tag{6.7.5}$$

onde usamos (6.6.12). As demais quantidades escalares envolvidas nesta última equação podem ser calculadas no referencial inercial que se revelar mais conveniente. Assim, calculando o segundo membro de (6.7.5) no referencial do centro de massa com a ajuda de (6.7.3), resulta

$$\mathbf{P}^{(F)^2} = 16m^2c^2 \ , \tag{6.7.6}$$

onde usamos (6.3.4b). Finalmente, no referencial do laboratório temos

$$\mathbf{P} = \left(\frac{E}{c}, \mathbf{p}\right) \ , \quad \mathbf{Q} = (mc, 0, 0, 0) \qquad \text{(no referencial do laboratório)} \ , \tag{6.7.7}$$

donde, por (6.3.4a),

$$\mathbf{P}\cdot\mathbf{Q} = \frac{E}{c}mc = mE \ . \tag{6.7.8}$$

Levando (6.7.6) e (6.7.8) em (6.7.5) obtém-se imediatamente

$$E = 7\,m\,c^2 \tag{6.7.9}$$

para a energia limiar da reação (6.7.2). Para possuir esta energia, o próton incidente deve ter uma velocidade igual a $0,99c$. ■

No caso especial de um fóton, sua energia é $h\nu$ e seu momento linear tem magnitude $h\nu/c$, onde h é a constante de Planck e ν é a freqüência. Assim, o quadrimomento de um fóton é

$$\mathbf{P} = \left(\frac{E}{c}, \mathbf{p}\right) = \frac{h\nu}{c}(1, \hat{\mathbf{n}}) \tag{6.7.10}$$

onde $\hat{\mathbf{n}}$ é o vetor unitário ao longo da direção de propagação. É claro que

$$\mathbf{P}^2 = 0 \ , \tag{6.7.11}$$

e se costuma dizer que $\mathbf{P}$ é um vetor *nulo*.

■ **Exemplo 6.7.2.** Interpretando o espalhamento de raios X por uma folha de grafite como provocado pela colisão de um fóton com um elétron, A. H. Compton mostrou, em 1923, que

$$\lambda' = \lambda + \frac{2h}{mc}\,\mathrm{sen}^2\,\frac{\theta}{2} \ . \tag{6.7.12}$$

Deduzir a fórmula de Compton, onde λ é o comprimento de onda do raio X incidente, λ' o do emergente, m é a massa do elétron e θ é o ângulo entre os raios X incidente e emergente.

Solução. Na Fig. 6.6 um fóton de freqüência ν colide com um elétron em repouso e é defletido de um ângulo θ ao mesmo tempo que sua freqüência ν é diminuída para ν', ao passo que o ângulo de recuo do elétron é θ'. Sejam $\mathbf{P}$ e $\mathbf{P}'$ os quadrimomentos inicial e final do fóton, com $\mathbf{Q}$ e $\mathbf{Q}'$ os do elétron. A conservação do quadrimomento requer

$$\mathbf{P} + \mathbf{Q} = \mathbf{P}' + \mathbf{Q}' \ . \tag{6.7.13}$$

Como não queremos nenhuma informação sobre o recuo do elétron, revela-se conveniente eliminar o quadrivetor indesejável $\mathbf{Q}'$, isolando-o em (6.7.13) e quadrando a equação resultante:

$$\mathbf{Q}'^2 = (\mathbf{P} + \mathbf{Q} - \mathbf{P}')^2 \ . \tag{6.7.14}$$

Como $\mathbf{Q}^2 = \mathbf{Q}'^2 = m^2c^2$ e $\mathbf{P}^2 = \mathbf{P}'^2 = 0$, ficamos com

$$\mathbf{P}\cdot\mathbf{P}' = \mathbf{Q}\cdot(\mathbf{P} - \mathbf{P}') \ . \tag{6.7.15}$$

Calculando o segundo membro desta última equação no referencial do laboratório, no qual $\mathbf{Q} = (mc, 0, 0, 0)$, encontramos $\mathbf{Q}\cdot\mathbf{P} = mh\nu$ e $\mathbf{Q}\cdot\mathbf{P}' = mh\nu'$, ao passo que

$$\mathbf{P}\cdot\mathbf{P}' = \frac{h\nu}{c}\frac{h\nu'}{c}(1 - \hat{\mathbf{n}}\cdot\hat{\mathbf{n}}') = \frac{h^2\nu\nu'}{c^2}(1 - \cos\,\theta) \ . \tag{6.7.16}$$

Levando estes últimos resultados em (6.7.15) obtemos, finalmente,

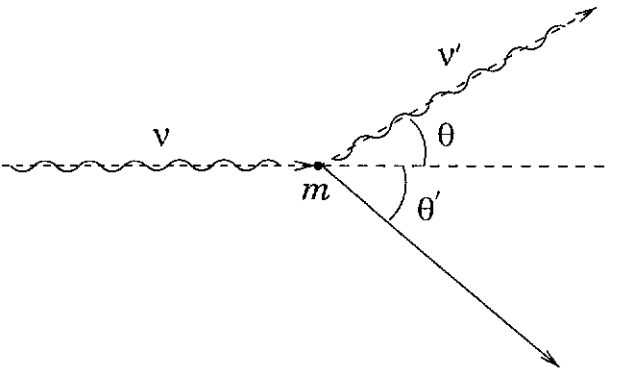


Figura 6.6: Efeito Compton.

$$\frac{h\nu\nu'}{c^2}(1-\cos\theta) = m(\nu-\nu') \ . \tag{6.7.17}$$

É fácil comprovar que, em termos do semi-ângulo $\theta/2$ e do comprimento de onda $\lambda = c/\nu$, esta equação torna-se idêntica à fórmula de Compton (6.7.12). ■

6.8 Dinâmica Relativística na Forma Lagrangiana

No caso de uma partícula sujeita a uma força $\mathbf{F}$ conservativa, isto é,

$$\mathbf{F} = -\boldsymbol{\nabla} V(\mathbf{r}) \ , \tag{6.8.1}$$

a equação de movimento relativística

$$\frac{d}{dt}\Big(\frac{m\,\mathbf{v}}{\sqrt{1-v^2/c^2}}\Big) = -\boldsymbol{\nabla} V(\mathbf{r}) \tag{6.8.2}$$

pode ser posta na forma lagrangiana. Para tanto basta notar que

$$\frac{m\,\dot{x}}{\sqrt{1-(\dot{x}^2+\dot{y}^2+\dot{z}^2)/c^2}} = \frac{\partial}{\partial \dot{x}}\Big[-mc^2\sqrt{1-(\dot{x}^2+\dot{y}^2+\dot{z}^2)/c^2}\Big] \ , \tag{6.8.3}$$

o que nos leva a concluir que a lagrangiana

$$L = -mc^2\sqrt{1-v^2/c^2} - V(\mathbf{r}) \tag{6.8.4}$$

permite escrever as componentes de (6.8.2) na forma

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_k}\Big) - \frac{\partial L}{\partial q_k} = 0 \quad , \quad k = 1,2,3 \ , \tag{6.8.5}$$

com $q_1 = x, q_2 = y, q_3 = z$. Note que a lagrangiana (6.8.4) *não* é da forma $T - V$.

Contudo, se V não depende explicitamente do tempo, uma constante de movimento de importância fundamental é a integral de Jacobi

$$h = \dot{x}\frac{\partial L}{\partial \dot{x}} + \dot{y}\frac{\partial L}{\partial \dot{y}} + \dot{z}\frac{\partial L}{\partial \dot{z}} - L = \frac{mc^2}{\sqrt{1 - v^2/c^2}} + V \ , \tag{6.8.6}$$

que coincide com a energia total da partícula.

A lagrangiana (6.8.4) difere de sua contraparte não-relativística apenas no termo cinético, o que nos induz a escrever a seguinte lagrangiana relativística para uma partícula num campo eletromagnético externo:

$$L = -mc^2\sqrt{1 - v^2/c^2} - e\phi + \frac{e}{c}\mathbf{v}\cdot\mathbf{A} \tag{6.8.7}$$

onde ϕ e $\mathbf{A}$ são os potenciais do campo eletromagnético. Se ϕ e $\mathbf{A}$ não dependem explicitamente do tempo, a integral de Jacobi coincide com a energia total:

$$h = E_{\text{tot}} = \frac{mc^2}{\sqrt{1 - v^2/c^2}} + e\phi \ . \tag{6.8.8}$$

A formulação lagrangiana que acabamos de discutir não é covariante, pois gera apenas a equação de movimento relativística (6.6.17), que se refere a um sistema inercial particular. Uma formulação lagrangiana covariante deve permitir escrever a equação de movimento covariante

$$\frac{d}{d\tau}(mU^\mu) = \mathcal{F}^\mu \ , \tag{6.8.9}$$

na forma

$$\frac{d}{d\tau}\left(\frac{\partial L}{\partial U_\mu}\right) - \frac{\partial L}{\partial x_\mu} = 0 \ , \tag{6.8.10}$$

onde x^μ e U^μ são tratadas como coordenadas e velocidades generalizadas, com τ desempenhando o papel do tempo. A lagrangiana L deve ser um escalar para que $\partial L/\partial x_\mu$ e $\partial L/\partial U_\mu$ sejam quadrivetores. No caso de um campo eletromagnético externo, uma lagrangiana simples que dá conta do recado é

$$L = \frac{m}{2}U^\alpha U_\alpha + \frac{e}{c}U_\alpha A^\alpha \ . \tag{6.8.11}$$

Notemos primeiro que

$$\frac{\partial}{\partial U_\mu}(U^\alpha U_\alpha) = \frac{\partial}{\partial U_\mu}(g^{\alpha\beta}U_\alpha U_\beta) = g^{\alpha\beta}(\delta^\mu_\alpha U_\beta + U_\alpha\delta^\mu_\beta) = \delta^\mu_\alpha U^\alpha + \delta^\mu_\beta U^\beta = 2U^\mu \ , \tag{6.8.12}$$

donde

$$\frac{\partial L}{\partial U_\mu} = mU^\mu + \frac{e}{c}A^\mu \ . \tag{6.8.13}$$

Por outro lado,

$$\frac{d}{d\tau}A^\mu = \frac{\partial A^\mu}{\partial x_\alpha}\frac{dx_\alpha}{d\tau} = U_\alpha \partial^\alpha A^\mu \ . \tag{6.8.14}$$

Assim, as equações de Lagrange (6.8.10) geradas pela lagrangiana (6.8.11) são

$$m\frac{dU^\mu}{d\tau} + \frac{e}{c}U_\alpha \partial^\alpha A^\mu - \frac{e}{c}U_\alpha \frac{\partial A^\alpha}{\partial x_\mu} = 0 \ , \tag{6.8.15}$$

que coincidem com (6.6.24).

Uma possível objeção à lagrangiana (6.8.11) é que ela só gera a equação de movimento covariante correta se tratarmos as quatro componentes da quadrivelocidade como independentes, o que obviamente não é verdadeiro por causa do vínculo (6.6.7). Pode-se contra-argumentar, no entanto, que além de produzir a equação de movimento correta, a própria equação de movimento gerada pela lagrangiana (6.8.11) contém o vínculo (6.6.7). De fato, de (6.8.15) deduz-se $U_\mu dU^\mu/d\tau = (e/mc)F^{\mu\nu}U_\mu U_\nu = 0$ porque $F^{\mu\nu}$ é anti-simétrico (ver Problema 6.6). Portanto, $U_\mu U^\mu =$ constante e a constante pode ser escolhida igual a c^2. Isto justifica o emprego da lagrangiana (6.8.11) tratando as componentes da quadrivelocidade como se fossem todas independentes.

6.9 Ação à Distância na Relatividade Especial

Na Seção 6.6 vimos como fazer a formulação covariante da dinâmica de uma partícula. O passo seguinte seria desenvolver equações de movimento manifestamente covariantes para duas ou mais partículas em interação. Uma das características da mecânica newtoniana é a ação à distância, tal como a expressa pela lei da gravitação universal. A conservação do momento linear total é uma conseqüência direta da terceira lei de Newton. Se a ação e a reação ocorrem em pontos distantes, essa lei não pode ser transplantada diretamente para a relatividade especial por causa da ambigüidade da simultaneidade: só é possível uma terceira lei de Newton covariante no caso de interação por contato. É evidente que a ação à distância nos moldes estritamente newtonianos é incompatível com a teoria especial da relatividade, porque envolve a propagação instantânea, isto é, com velocidade infinita, da interação entre as partículas. Para contornar essa dificuldade, poder-se-ia conceber uma ação à distância retardada, de tal modo que qualquer modificação no estado de uma partícula somente seria sentida por outra depois de transcorrido um tempo igual ao necessário para a luz cruzar a distância entre elas. Nesse caso, a modificação do momento linear de uma partícula não se refletiria na mudança imediata do momento linear das outras, mas teria que haver um momento linear em trânsito de uma partícula para as outras a fim de assegurar a conservação do momento linear total. Isto indica que, numa teoria relativística, as interações entre partículas têm que ser mediadas por campos locais dotados de propriedades mecânicas, tais como energia e momento linear. Embora isto seja intuitivamente claro, uma demonstração formal da impossibilidade de ação à distância na teoria especial da relatividade tem um considerável valor instrutivo.

Considere um sistema de duas partículas com linhas de universo descritas pelos quadrivetores $x^\mu_{(1)}(\tau_1)$ e $x^\mu_{(2)}(\tau_2)$, onde τ_1 e τ_2 são os respectivos tempos próprios. O quadrimomento total do sistema no instante t de um referencial inercial $\mathcal{K}$ é

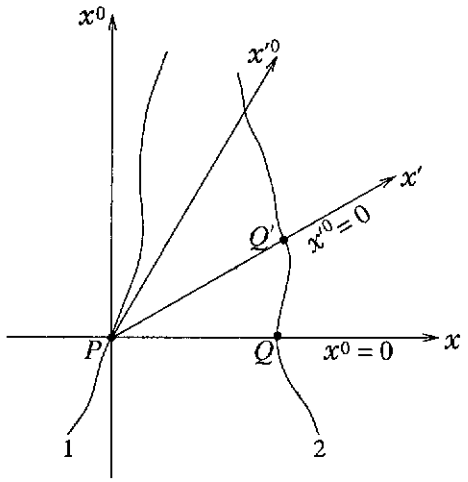


Figura 6.7: Linhas de universo de duas partículas em interação.

$$\mathbf{P}(t) = \mathbf{P}_{(1)}(\tau_1) + \mathbf{P}_{(2)}(\tau_2) \ , \tag{6.9.1}$$

onde os tempos próprios τ_1 e τ_2 correspondem ao mesmo instante t nas linhas de universo das partículas. Uma teoria relativística razoável de duas partículas em interação via ação à distância deve possuir as seguintes propriedades: (a) o quadrimomento total é conservado em todos os referenciais da teoria especial da relatividade; (b) assintoticamente, isto é, para $\tau_i \to \infty$ ou $\tau_i \to -\infty$ as linhas de universo são retas, isto é, as partículas são livres. Esta última hipótese tem por principal objetivo possibilitar a descrição de um processo de espalhamento, em que inicialmente as partículas estão muito afastadas, aproximam-se e interagem durante um curto intervalo de tempo, e depois afastam-se indefinidamente seguindo trajetórias retilíneas desviadas de suas direções originais. Esta condição assintótica também assegura que o quadrimomento total se transforma como um quadrivetor. Ora, direis, mas isto não é óbvio? Parece, mas não é, porque em (6.9.1) os eventos correspondentes aos tempos próprios τ_1 e τ_2 são simultâneos no referencial $\mathcal{K}$ mas não são simultâneos em nenhum outro referencial inercial em movimento relativamente a $\mathcal{K}$. No entanto, assintoticamente esses quadrimomentos tornam-se constantes e podem ser calculados em qualquer instante. Assim, podemos considerá-los calculados simultaneamente em qualquer sistema inercial, de modo que o quadrimomento total transforma-se como um quadrivetor. Mas, como a lei de transformação entre referenciais inerciais tem que ser a mesma em qualquer instante, se o quadrimomento total transforma-se como um quadrivetor na região assintótica a mesma lei de transformação aplica-se em qualquer instante, isto é:

$$P'^{\mu} = \Lambda^{\mu}{}_{\nu} P^{\nu} \ . \tag{6.9.2}$$

Consideremos dois sistemas de referência inerciais $\mathcal{K}$ e $\mathcal{K}'$ com origem comum situada sobre a linha de universo da primeira partícula, conforme ilustrado na Fig. 6.7. Pela hipótese de conservação do quadrimomento total, cada lado da Eq.(6.9.2) é constante e

pode ser calculado em qualquer instante t ou t' que se desejar. Explorando essa liberdade, calculemos P'^{μ} no instante $t' = 0$ e P^{μ} no instante $t = 0$ para obter

$$P'^{\mu}_{(1)}\big|_P + P'^{\mu}_{(2)}\big|_{Q'} = \Lambda^{\mu}{}_{\nu}\,(P^{\nu}_{(1)}\big|_P + P^{\nu}_{(2)}\big|_Q) \ . \tag{6.9.3}$$

Mas, nos pontos P e Q, a transformação de Lorentz do quadrimomento fornece

$$P'^{\mu}_{(1)}\big|_P = \Lambda^{\mu}{}_{\nu}\,P^{\nu}_{(1)}\big|_P \quad , \quad P'^{\mu}_{(2)}\big|_{Q'} = \Lambda^{\mu}{}_{\nu}\,P^{\nu}_{(2)}\big|_{Q'} \ . \tag{6.9.4}$$

Combinando (6.9.3) e (6.9.4) resulta

$$\Lambda^{\mu}{}_{\nu}\,P^{\nu}_{(2)}\big|_{Q'} = \Lambda^{\mu}{}_{\nu}\,P^{\nu}_{(2)}\big|_Q \ . \tag{6.9.5}$$

Como a matriz $\mathbf{\Lambda}$ é inversível, segue-se que

$$P^{\mu}_{(2)}\big|_{Q'} = P^{\mu}_{(2)}\big|_Q \ , \tag{6.9.6}$$

o que significa que o quadrimomento da partícula número dois é o mesmo nos pontos Q e Q' de sua linha de universo. Como Q e Q' são eventos arbitrários e podemos intercambiar os papéis das partículas, demonstramos que o quadrimomento total será conservado se e somente se o quadrimomento de cada partícula permanecer constante, isto é, se não houver interação entre as partículas.

Este "teorema da ausência de interação" (Van Dam & Wigner 1966; Mann 1974; Ohanian 1976) estabelece de forma inequívoca que a implementação da lei da conservação da energia e do momento linear de um sistema de partículas em interação requer a introdução de uma outra entidade portadora de energia e momento linear, a saber, um campo mediador da interação. A interação entre duas partículas afastadas deve ser interpretada como uma ação por contato de uma partícula sobre o campo, seguida de propagação do campo e, finalmente, ação por contato do campo sobre a outra partícula. A quantidade que se conserva é a soma dos quadrimomentos das partículas acrescida do quadrimomento do campo.

PROBLEMAS

6.1. Dois fótons viajam ao longo do eixo x do referencial inercial $\mathcal{K}$ mantendo uma distância constante L entre si. Prove que em $\mathcal{K}'$ a distância entre esses fótons é dada por $L(c+v)^{1/2}/(c-v)^{1/2}$.

6.2. (i) Mostre que a transformação de Lorentz padrão (6.1.1) pode ser escrita na forma

$$x' = x\cosh\phi - ct\,\mathrm{senh}\,\phi \quad , \quad ct' = -x\,\mathrm{senh}\,\phi + ct\cosh\phi \ ,$$

onde $\tanh\phi = v/c$. (As relações $\cosh\phi = \cos i\phi$ e $i\,\mathrm{senh}\,\phi = \mathrm{sen}\,i\phi$ permitem converter qualquer identidade trigonométrica numa identidade de funções hiperbólicas.) Note que, formalmente, isto é uma "rotação" em x e ict de um "ângulo" $i\phi$; como tal, ela preserva $x^2 + (ict)^2$. (ii) Deduza a seguinte forma útil da transformação de Lorentz:

$$ct' + x' = e^{-\phi}(ct + x) \quad , \quad ct' - x' = e^{\phi}(ct - x) \ .$$

Se S'' é um referencial que se move com velocidade u ao longo da direção x' de S', use esta forma da transformação de Lorentz para provar que a velocidade de S'' em relação a S é $(u+v)/(1+uv/c^2)$.

6.3. Se $\mathbf{v}'$ é a velocidade de uma partícula vista do referencial inercial $\mathcal{K}'$ que se move em relação ao referencial $\mathcal{K}$ com velocidade $\mathbf{u}$, prove que a velocidade da partícula em relação a $\mathcal{K}$ é

$$\mathbf{v} = \frac{\mathbf{v}' + (\gamma - 1)(\mathbf{u}\cdot\mathbf{v}'/u^2)\mathbf{u} + \gamma\mathbf{u}}{\gamma\left(1 + \mathbf{u}\cdot\mathbf{v}'/c^2\right)} \ ,$$

onde $\gamma = (1 - u^2/c^2)^{-1/2}$. Note que, em geral, esta expressão não é simétrica em $\mathbf{u}$ e $\mathbf{v}'$. A que se reduz esta equação quando $\mathbf{v}'$ e $\mathbf{u}$ são paralelas?

6.4. Dois eventos A e B têm uma separação do tipo tempo. Considere uma linha de universo reta e linhas de universo curvas conectando A e B. Mostre que o tempo próprio entre A e B,

$$\Delta\tau = \int_A^B d\tau = \int_A^B \sqrt{dx^\mu dx_\mu}$$

é máximo quando calculado ao longo da linha de universo reta. Assim, a linha de universo de uma partícula livre pode ser caracterizada como a curva no espaço-tempo de máximo tempo próprio. Sugestão: use um sistema de referência em que a partícula movendo-se ao longo da linha de universo reta permaneça em repouso espacial.

6.5. Uma transformação de Lorentz infinitesimal é da forma $\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu$, onde os $\omega^\mu{}_\nu$ são quantidades infinitesimais. Usando (6.1.19) e desprezando infinitésimos de ordem superior à primeira, prove que $\omega_{\mu\nu} = -\omega_{\nu\mu}$.

6.6. (i) Prove que um tensor $A^{\mu\nu}$ é simétrico (anti-simétrico) na forma contravariante se e somente se ele é simétrico (anti-simétrico) na forma covariante $A_{\mu\nu}$. (ii) Se $A^{\mu\nu}$ é um tensor simétrico e $B^{\mu\nu}$ é um tensor anti-simétrico, prove que $A^{\mu\nu}B_{\mu\nu} = 0$.

6.7. Mostre que um fóton não pode desintegrar-se espontaneamente num par elétron-pósitron.

6.8. (i) Prove que na mecânica relativística também vale a equação $d\mathbf{L}/dt = \mathbf{N}$, onde $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ é o momento angular e $\mathbf{N} = \mathbf{r} \times \mathbf{F}$ é o torque ($\mathbf{p}$ é o momento linear relativístico). (ii) Se $L^{\mu\nu} = x^\mu P^\nu - x^\nu P^\mu$ é o tensor momento angular e $N^{\mu\nu} = x^\mu \mathcal{F}^\nu - x^\nu \mathcal{F}^\mu$ é o tensor torque, demonstre que $dL^{\mu\nu}/d\tau = N^{\mu\nu}$ e comprove que a parte espacial desta equação corresponde ao resultado do item (i).

6.9. (i) Se A^μ é um vetor do tipo tempo, prove que existe um referencial inercial no qual ele só possui componente temporal. Sugestão: uma rotação dos eixos permite eliminar duas das componentes espaciais do vetor; prove que existe uma transformação de Lorentz que elimina a componente espacial remanescente. (ii) Prove que se um quadrivetor é ortogonal a um quadrivetor do tipo tempo, então ele é do tipo espaço.

6.10. Uma partícula de massa M decai a partir do repouso numa partícula de massa m e num fóton. Encontre a energia de cada um dos produtos finais.

6.11. Prove que uma partícula de massa não-nula não pode desintegrar-se num único fóton. Que lei de conservação é necessariamente violada?

6.12. No referencial do laboratório um fóton colide com um próton em repouso. Qual é a freqüência mínima do fóton para que os produtos da colisão sejam um nêutron e um méson pi? Denote por m a massa do próton ou nêutron, supostas iguais, e por m_π a massa do píon.

6.13. Usando a lei de transformação da quadriaceleração $\mathcal{A}^\mu$, mostre que as componentes da aceleração tridimensional $\mathbf{a}'$ no referencial $\mathcal{K}'$ em que a partícula se encontra instantaneamente em repouso são

$$a'_x = \frac{a_x}{(1 - v^2/c^2)^{3/2}} \quad , \quad a'_y = \frac{a_y}{1 - v^2/c^2} \quad , \quad a'_z = \frac{a_z}{1 - v^2/c^2} \ ,$$

onde os eixos x e x' foram escolhidos ao longo da velocidade $\mathbf{v}$ da partícula em relação ao referencial $\mathcal{K}$. Mostre que a aceleração própria no movimento hiperbólico (Exemplo 6.6.1) é constante e igual a F/m.

6.14. Um píon com energia de 1200 MeV incide sobre um próton em repouso, produzindo vários píons segundo a reação $\pi + p \to p + n\pi$. Qual é o número máximo n de píons que podem ser produzidos na reação? A energia de repouso do próton é 938 MeV e a do píon 140 MeV.

6.15. Uma partícula de massa de repouso m, com velocidade $v = 4c/5$, colide inelasticamente com uma partícula idêntica em repouso. (i) Qual é a velocidade da partícula composta? (ii) Qual é a sua massa de repouso?

6.16. Uma partícula de massa M, em repouso, decai em duas partículas de massas m_1 e m_2. Mostre que a energia da primeira partícula no referencial de repouso da partícula que se desintegrou é $E_1 = (M^2 + m_1^2 - m_2^2)/2M$.

6.17. O "dual" do campo eletromagnético é o tensor anti-simétrico $\tilde{F}^{\mu\nu}$ definido por

$$\tilde{F}^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\rho\lambda} F_{\rho\lambda} \ ,$$

onde $\epsilon^{\mu\nu\rho\lambda}$ é a generalização quadridimensional do símbolo de Levi-Civita introduzido no Apêndice A: $\epsilon^{\mu\nu\rho\lambda} = \pm 1$ conforme $(\mu, \nu, \rho, \lambda)$ seja uma permutação par ou ímpar de $(0, 1, 2, 3)$, e $\epsilon^{\mu\nu\rho\lambda} = 0$ se houver repetição de índices. (i) Mostre que as componentes de $\tilde{F}^{\mu\nu}$ podem ser obtidas simplesmente trocando $\mathbf{E}$ por $\mathbf{B}$ e $\mathbf{B}$ por $-\mathbf{E}$ em $F^{\mu\nu}$. (ii) Prove que as equações de Maxwell homogêneas podem ser escritas na forma

$$\partial_\nu \tilde{F}^{\mu\nu} = 0 \ .$$

(iii) Exprima os escalares de Lorentz $F^{\mu\nu}F_{\mu\nu}$, $\tilde{F}^{\mu\nu}\tilde{F}_{\mu\nu}$ e $\tilde{F}^{\mu\nu}F_{\mu\nu}$ em termos de $\mathbf{E}$ e $\mathbf{B}$. (iv) Prove que se $\mathbf{E}$ e $\mathbf{B}$ são perpendiculares num dado referencial inercial, eles são perpendiculares em todos os referenciais inerciais.

6.18. O tensor de energia-momento do campo eletromagnético é definido por

$$T^{\mu\nu} = \frac{1}{4\pi}\left[F^{\alpha\mu}F^{\nu}{}_{\alpha} + \frac{1}{4}\, g^{\mu\nu}F^{\alpha\beta}F_{\alpha\beta}\right] \ .$$

Mostre que:

$$T^{00} = \frac{1}{8\pi}\left(E^2 + B^2\right) = \text{densidade de energia} \ ;$$

para $k = 1, 2, 3$

$$T^{0k} = \frac{1}{4\pi}\left(\mathbf{E} \times \mathbf{B}\right)_k = \text{vetor de Poynting} \ .$$

Mostre, ainda, que o traço de $T^{\mu\nu}$ é zero, isto é, $T^{\mu}{}_{\mu} = 0$.

6.19. Considere uma possível modificação da teoria do campo gravitacional visando torná-la relativisticamente covariante: a expressão usual da força $\mathbf{F} = -m\nabla\phi$ é substituída por $\mathcal{F}_\mu = m\partial_\mu \phi$, onde m é a massa da partícula e ϕ é o potencial gravitacional; na equação de movimento covariante $d(mU^\mu)/d\tau = \mathcal{F}^\mu$ a massa da partícula é suposta variável. Deduza a equação $c^2 dm/d\tau = m d\phi/d\tau$ e, como conseqüência, $m = m_0 \exp(\phi/c^2)$, com m_0 constante. Segundo tal teoria, portanto, a massa de cada partícula dependeria do potencial gravitacional no ponto em que ela se encontra. Esta teoria foi proposta por G. Nordström em 1912, tendo sido abandonada logo em seguida (Pais 1982).

6.20. O vínculo $U^\mu U_\mu = c^2$ pode ser incorporado à dinâmica lagrangiana da partícula livre relativística por meio de um multiplicador de Lagrange, tratado como um grau de liberdade adicional do sistema. Considere, para uma partícula livre, a lagrangiana

$$L(x^\mu, U^\mu, \lambda, \dot{\lambda}) = \frac{\lambda}{2}\left(U^\mu U_\mu - c^2\right) \ .$$

Mostre que

$$\frac{d}{d\tau}\left(\frac{\partial L}{\partial U^\alpha}\right) - \frac{\partial L}{\partial U^\alpha} = 0 \quad \Longrightarrow \quad \lambda \frac{dU_\alpha}{d\tau} + \dot{\lambda}\, U_\alpha = 0$$

e

$$\frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{\lambda}}\right) - \frac{\partial L}{\partial \lambda} = 0 \quad \Longrightarrow \quad U^\mu U_\mu - c^2 = 0 \ .$$

Destas equações deduza que $\dot{\lambda} = 0$, isto é, $\lambda =$ constante. Conclua, finalmente, que as equações de movimento acima equivalem ao par

$$\frac{dU^\alpha}{d\tau} = 0 \quad , \quad U^\alpha U_\alpha = c^2 \ ,$$

que são as equações corretas para a partícula livre, incluindo o vínculo satisfeito pela quadrivelocidade.

Capítulo 7

DINÂMICA HAMILTONIANA

As equações da mecânica analítica têm um significado que excede em muito o da mecânica newtoniana.

Albert Einstein

Na formulação lagrangiana o movimento de um sistema mecânico com n graus de liberdade é regido por n equações diferenciais ordinárias de segunda ordem no tempo. A dinâmica de Hamilton consiste em substituir as n equações de Lagrange por um certo conjunto equivalente de $2n$ equações diferenciais ordinárias de primeira ordem. De modo geral, as equações de Hamilton para problemas específicos não são mais fáceis de resolver do que as equações de Lagrange. A importância do formalismo hamiltoniano reside em fornecer um método poderoso, geral e flexível para a investigação de questões *estruturais* da mecânica, e, sobretudo, em servir de fundamento para a mecânica quântica e a mecânica estatística. Neste e nos próximos capítulos discutiremos como efetuar a transição do formalismo de Lagrange para o de Hamilton, bem como as inúmeras vantagens formais inerentes à descrição hamiltoniana da mecânica clássica.

7.1 As Equações Canônicas de Hamilton

As equações de Lagrange para um sistema com n graus de liberdade constituem um conjunto de n equações diferenciais ordinárias de segunda ordem no tempo para as coordenadas generalizadas $q_1(t), \ldots, q_n(t)$. O movimento do sistema é univocamente determinado desde que $2n$ condições iniciais sejam especificadas, a saber, os valores de todos os qs e $\dot{q}$s num instante particular t_0. O movimento pode ser representado geometricamente por uma curva traçada no espaço de configuração, cujas coordenadas são $q_1, \ldots, q_n$.

Na formulação introduzida por Hamilton o quadro é diferente: as equações de movimento são $2n$ equações diferenciais ordinárias de primeira ordem no tempo para $2n$ variáveis independentes. O movimento pode ser representado por uma curva traçada no *espaço de fase*, cujas coordenadas são as referidas variáveis independentes. Diferentemente do espaço de configuração, no qual um ponto define apenas a configuração (posições das partículas) do sistema num dado instante, um ponto do espaço de fase determina o

estado do sistema, isto é, sua configuração (posições das partículas) e a taxa de variação temporal desta configuração (velocidades das partículas) num dado instante.

Naturalmente, as equações de Lagrange podem ser trivialmente substituídas por um sistema equivalente de primeira ordem, com o número de equações duplicado, meramente introduzindo as variáveis $s_i = \dot{q}_i$, $i = 1, \ldots, n$ e tratando $q_1, \ldots, q_n, s_1, \ldots, s_n$ como $2n$ variáveis independentes. As equações de movimento seriam

$$\dot{q}_i = s_i \quad , \quad \frac{d}{dt}\left(\frac{\partial L}{\partial s_i}\right) - \frac{\partial L}{\partial q_i} = 0 \quad , \quad i = 1, \ldots, n \ , \tag{7.1.1}$$

onde $L(q, s, t)$ é a lagrangiana do sistema. Estas equações, no entanto, envolvem q_i e s_i de forma assimétrica e não são particularmente úteis.

Momentos Canônicos e Equações de Hamilton

William Rowan Hamilton mostrou, em 1835, que a duplicação simétrica do número de variáveis independentes é conseguida graças à descrição da dinâmica por intermédio das $2n$ quantidades $q_1, \ldots, q_n, p_1, \ldots, p_n$ onde p_i é o *momento canônico conjugado* a q_i, definido por

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \quad , \quad i = 1, \ldots, n \ . \tag{7.1.2}$$

Suporemos que a *matriz hessiana* $\mathbf{W} = (W_{ij})$ com elementos

$$W_{ij} = \frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j} \tag{7.1.3}$$

é não-singular, isto é,

$$\det \mathbf{W} \neq 0 \ , \tag{7.1.4}$$

pois neste caso o teorema da função implícita (Spivak 1965; Protter & Morrey Jr. 1985) garante que as Eqs.(7.1.2) podem ser resolvidas para as velocidades generalizadas. Assim, a descrição hamiltoniana envolve a substituição das variáveis $(q, \dot{q})$ por (q, p) em todas as grandezas mecânicas, e a introdução de uma função $H(q, p, t)$ em lugar da lagrangiana $L(q, \dot{q}, t)$ para gerar a dinâmica. Tal mudança de descrição realiza-se mediante uma *transformação de Legendre*,[1] que no presente contexto consiste na substituição das velocidades generalizadas pelos momentos canônicos como variáveis básicas e na introdução da *função de Hamilton* ou, simplesmente, *hamiltoniana* $H(q, p, t)$ definida por

$$H(q, p, t) = \sum_{i=1}^{n} \dot{q}_i \, p_i \ - \ L(q, \dot{q}, t) \ . \tag{7.1.5}$$

Supõe-se que no segundo membro desta equação as velocidades estejam expressas na forma $\dot{q}_i = f_i(q, p, t)$ resultante da resolução das n equações (7.1.2) para as n velocidades generalizadas.

[1] As transformações de Legendre são largamente utilizadas na termodinâmica e também possuem um significado geométrico interessante (Callen 1960; Arnold 1976) que não poderemos discutir aqui.

As conseqüências mais imediatas da introdução da função H podem ser deduzidas tomando-se a diferencial da Eq.(7.1.5):

$$dH = \sum_{i=1}^{n} (\dot{q}_i dp_i + p_i d\dot{q}_i) - \left\{ \sum_{i=1}^{n} \left(\frac{\partial L}{\partial q_i} dq_i + \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i \right) + \frac{\partial L}{\partial t} dt \right\} . \tag{7.1.6}$$

Em virtude da definição (7.1.2) dos momentos canônicos e das equações de Lagrange, esta última equação reduz-se a

$$dH = \sum_{i=1}^{n} (\dot{q}_i dp_i - \dot{p}_i dq_i) - \frac{\partial L}{\partial t} dt , \tag{7.1.7}$$

indicando que, de fato, H só depende dos qs e ps. Por outro lado,

$$dH = \sum_{i=1}^{n} \left(\frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial p_i} dp_i \right) + \frac{\partial H}{\partial t} dt . \tag{7.1.8}$$

Comparando estas duas últimas equações resulta

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad , \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \quad , \quad i = 1, \ldots, n , \tag{7.1.9}$$

e, como subproduto,

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} . \tag{7.1.10}$$

As equações (7.1.9) são conhecidas como *equações de Hamilton* ou *equações canônicas de Hamilton*, e formam um conjunto de $2n$ equacões diferenciais de primeira ordem equivalente ao sistema de n equações de segunda ordem de Lagrange. As quantidades (q, p) são chamadas de *variáveis canônicas*[2] e o espaço cartesiano de $2n$ dimensões cujos pontos são representados pelas $2n$-uplas $(q, p) \equiv (q_1, \ldots, q_n, p_1, \ldots, p_n)$ é chamado de *espaço de fase*. Um ponto no espaço de fase define o estado do sistema (posições e velocidades das partículas) num dado instante. Um teorema de existência e unicidade garante que a especificação do estado do sistema num instante t_0 determina uma única solução $(q(t), p(t))$ para as equações de Hamilton. Assim, por cada ponto do espaço de fase passa uma única trajetória dinâmica $(q(t), p(t))$, e duas trajetórias dinâmicas distintas nunca se tocam. Quanto a (7.1.10), não se trata de uma equação de movimento, mas de uma importante relação entre as dependências temporais explícitas da lagrangiana e da hamiltoniana. A primeira metade das equações de Hamilton exprime as velocidades generalizadas em termos das variáveis canônicas, ou seja, elas são as inversas das equações (7.1.2) que definem os momentos canônicos p_i. Ressalte-se, no entanto, que, ao contrário da formulação lagrangiana, na qual vale *a priori* a conexão $\dot{q}_i = dq_i/dt$, na dinâmica de Hamilton não há qualquer conexão *a priori* entre as variáveis canônicas, isto é, os qs e ps são inteiramente independentes entre si. Por isso as duas metades das Eqs.(7.1.9) devem ser encaradas em pé de igualdade, constituindo o conjunto completo de equações de movimento do sistema.

Excetuados certos casos em que a hamiltoniana pode ser escrita diretamente, a construção das equações de Hamilton envolve os seguintes estágios:

[2] A palavra "canônica" é usada no sentido de "padrão".

(a) Escolhidas coordenadas generalizadas, constrói-se a lagrangiana $L(q, \dot{q}, t)$.
(b) As Eqs.(7.1.2) são resolvidas para as velocidades $\dot{q}_i$ como funções de (q, p, t).
(c) Constrói-se $H(q, p, t)$ substituindo-se em (7.1.5) os $\dot{q}$s obtidos no passo anterior.
(d) Uma vez obtida $H(q, p, t)$, as equações de movimento do sistema são (7.1.9).

Hamiltoniana e Energia Total

A lagrangiana usual é $L = T - V$. Se (i) T é função puramente quadrática das velocidades e (ii) V independe das velocidades, então $\sum_i \dot{q}_i \partial L/\partial \dot{q}_i = \sum_i \dot{q}_i \partial T/\partial \dot{q}_i = 2T$ pelo teorema de Euler das funções homogêneas. Em conseqüência,

$$H = T + V = E \ , \tag{7.1.11}$$

isto é, *a hamiltoniana é a energia total expressa como função das coordenadas e momentos.* As condições (i) e (ii) prevalecem na maioria esmagadora dos casos de interesse físico, de modo que a hamiltoniana possui o significado físico extremamente importante se ser a energia total do sistema na vasta maioria das situações relevantes. Além disso, essas condições são apenas suficientes, sendo possível que elas não sejam satisfeitas e, ainda assim, H coincida com a energia total (vide Exemplo 7.1.2).

■ **Exemplo 7.1.1.** Obter as equações de Hamilton para uma partícula num potencial central. **Solução.** Em coordenadas esféricas a lagrangiana escreve-se

$$L = T - V = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \operatorname{sen}^2\theta\, \dot{\varphi}^2) - V(r) \ , \tag{7.1.12}$$

donde

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r} \quad , \quad p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta} \quad , \quad p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = mr^2 \operatorname{sen}^2\theta\, \dot{\varphi} \ . \tag{7.1.13}$$

Resolvendo estas equações para as velocidades encontramos

$$\dot{r} = \frac{p_r}{m} \quad , \quad \dot{\theta} = \frac{p_\theta}{mr^2} \quad , \quad \dot{\varphi} = \frac{p_\varphi}{mr^2 \operatorname{sen}^2\theta} \ , \tag{7.1.14}$$

de sorte que

$$H = \dot{r}\, p_r + \dot{\theta}\, p_\theta + \dot{\varphi}\, p_\varphi - L = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\varphi^2}{r^2 \operatorname{sen}^2\theta}\right) + V(r) \ . \tag{7.1.15}$$

Esta hamiltoniana é igual à energia total da partícula e as equações de Hamilton são

$$\dot{r} = \frac{\partial H}{\partial p_r} = \frac{p_r}{m} \quad , \quad \dot{\theta} = \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{mr^2} \quad , \quad \dot{\varphi} = \frac{\partial H}{\partial p_\varphi} = \frac{p_\varphi}{mr^2 \operatorname{sen}^2\theta} \ , \tag{7.1.16a}$$

$$\dot{p}_r = -\frac{\partial H}{\partial r} = \frac{p_\theta^2}{mr^3} + \frac{p_\varphi^2}{mr^3 \operatorname{sen}^2\theta} - \frac{dV}{dr} \ ; \ \dot{p}_\theta = -\frac{\partial H}{\partial \theta} = \frac{p_\varphi^2 \cot\theta}{mr^2 \operatorname{sen}^2\theta} \ , \ \dot{p}_\varphi = -\frac{\partial H}{\partial \varphi} = 0 \ . \tag{7.1.16b}$$

Conforme uma observação geral já feita, as Eqs.(7.1.16a) são as inversas de (7.1.13), e com o seu uso as Eqs.(7.1.16b) tornam-se idênticas às equações de Lagrange oriundas da lagrangiana (7.1.12). ■

■ **Exemplo 7.1.2.** Construir a hamiltoniana e as equações de Hamilton para uma partícula num campo eletromagnético externo.

Solução. Em coordenadas cartesianas temos

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - e\phi(\mathbf{r}, t) + \frac{e}{c}\mathbf{v}\cdot\mathbf{A}(\mathbf{r}, t) \ , \tag{7.1.17}$$

donde

$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x} + \frac{e}{c}A_x \quad , \quad p_y = \frac{\partial L}{\partial \dot{y}} = m\dot{y} + \frac{e}{c}A_y \quad , \quad p_z = \frac{\partial L}{\partial \dot{z}} = m\dot{z} + \frac{e}{c}A_z \ . \tag{7.1.18}$$

Assim,

$$\mathbf{v} = \frac{1}{m}\Big(\mathbf{p} - \frac{e}{c}\mathbf{A}\Big) \tag{7.1.19}$$

e, conseqüentemente,

$$H = \mathbf{v}\cdot\mathbf{p} - L = \frac{1}{2m}\Big(\mathbf{p} - \frac{e}{c}\mathbf{A}\Big)^2 + e\phi \ . \tag{7.1.20}$$

Esta hamiltoniana é a energia total se ϕ e $\mathbf{A}$ não dependem explicitamente do tempo (campos $\mathbf{E}$ e $\mathbf{B}$ estáticos). As equações de Hamilton tomam a forma

$$\dot{\mathbf{r}} = \frac{\partial H}{\partial \mathbf{p}} = \frac{1}{m}\Big(\mathbf{p} - \frac{e}{c}\mathbf{A}\Big) \ , \tag{7.1.21a}$$

$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{r}} \equiv -\nabla H = \frac{e}{mc}\left[\Big(\mathbf{p} - \frac{e}{c}\mathbf{A}\Big)\cdot\nabla\,\mathbf{A} + \Big(\mathbf{p} - \frac{e}{c}\mathbf{A}\Big)\times(\nabla\times\mathbf{A})\right] - e\nabla\phi \ , \tag{7.1.21b}$$

onde usamos $\nabla(\mathbf{G}\cdot\mathbf{G}) = 2(\mathbf{G}\cdot\nabla)\mathbf{G} + 2\mathbf{G}\times(\nabla\times\mathbf{G})$ com $\mathbf{G} = \mathbf{p} - (e/c)\,\mathbf{A}$. ■

■ **Exercício 7.1.1.** Substituindo a primeira das equações de Hamilton (7.1.21) na segunda, mostre que elas são equivalentes a

$$m\ddot{\mathbf{r}} = e\Big(\mathbf{E} + \frac{\mathbf{v}}{c}\times\mathbf{B}\Big) \ , \tag{7.1.22}$$

que é a equação de movimento newtoniana. ■

7.2 Simetrias e Leis de Conservação

Conforme definimos na Seção 2.5, uma coordenada q_j é dita cíclica ou ignorável se ela não aparece na lagrangiana. Em decorrência da definição dos momentos canônicos e da hamiltoniana, q_j também não aparece em H. Portanto, segue-se das equações de Hamilton que

$$\dot{p}_j = -\frac{\partial H}{\partial q_j} = 0 \tag{7.2.1}$$

e o momento conjugado a q_j é constante de movimento. De forma ligeiramente modificada,

a conexão entre simetrias e leis de conservação, já estudada no formalismo lagrangiano, pode ser reformulada na linguagem hamiltoniana.

■ **Exercício 7.2.1.** Considere um sistema de partículas descrito, em coordenadas cartesianas, pela hamiltoniana $H(\mathbf{r}_1, \ldots, \mathbf{r}_N, \mathbf{p}_1, \ldots, \mathbf{p}_N)$. (i) Se H é invariante sob a translação infinitesimal $\delta\mathbf{r}_k = \epsilon\hat{\mathbf{n}}$, $k = 1, \ldots N$, prove que a componente do momento linear canônico total $\mathbf{P} = \sum_k \mathbf{p}_k$ ao longo da direção $\hat{\mathbf{n}}$ é constante de movimento. (ii) Se H é invariante sob a rotação infinitesimal $\delta\mathbf{r}_k = \delta\theta\hat{\mathbf{n}} \times \mathbf{r}_k$, $\delta\mathbf{p}_k = \delta\theta\hat{\mathbf{n}} \times \mathbf{p}_k$, prove que a componente do momento angular canônico total $\mathbf{L} = \sum_k \mathbf{r}_k \times \mathbf{p}_k$ ao longo do eixo de rotação é constante de movimento. ■

Na mecânica lagrangiana sabe-se que se $\partial L/\partial t = 0$ então a função energia h é uma constante de movimento. Por outro lado, a definição (7.1.5) mostra que, embora dependam de variáveis diferentes, H e h têm o mesmo valor. Assim sendo, deve haver um teorema de conservação correspondente para H. Com efeito, pelas equações de Hamilton,

$$\frac{dH}{dt} = \sum_i \left(\frac{\partial H}{\partial q_i}\dot{q}_i + \frac{\partial H}{\partial p_i}\dot{p}_i \right) + \frac{\partial H}{\partial t} = \sum_i \left(\frac{\partial H}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial p_i}\frac{\partial H}{\partial q_i} \right) + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t} \ . \quad (7.2.2)$$

Desta equação infere-se um resultado simples e importante.

Teorema 7.2.1. Se a hamiltoniana não depende explicitamente do tempo, isto é, $\partial H/\partial t = 0$, então H é constante de movimento.

Vale destacar que, conforme (7.1.10), H depende explicitamente do tempo se e somente se o mesmo acontece com L. Por outro lado, como já vimos, em quase todas as situações de interesse físico H não depende explicitamente do tempo e coincide com a energia total do sistema. Assim, o Teorema 7.2.1 efetivamente constitui a lei da conservação da energia na vasta maioria das circunstâncias. Em casos excepcionais, todavia, pode acontecer de H se conservar sem ser a energia total ou, inversamente, de H ser a energia total sem conservar-se.

■ **Exemplo 7.2.1.** Reconsiderando o Exemplo 1.5.2, obter hamiltoniana e estudar sua conservação, assim como a da energia.

Solução. Como vimos no Exemplo 1.5.2, a lagrangiana do sistema é

$$L = T = \frac{m}{2}(\dot{r}^2 + \omega^2 r^2) \ , \quad (7.2.3)$$

que coincide com a energia cinética da conta. Neste caso, $p_r = m\dot{r}$ e

$$H = \frac{p_r^2}{2m} - \frac{m\omega^2}{2}r^2 \ , \quad (7.2.4)$$

que *não* é a energia total (puramente cinética) da conta. Entretanto, H é constante de movimento porque $\partial H/\partial t = 0$. Por outro lado, a energia total

$$E = T = \frac{p_r^2}{2m} + \frac{m\omega^2}{2}r^2 \quad (7.2.5)$$

não se conserva porque a força de vínculo realiza trabalho por ocasião de um deslocamento

real da partícula (comprove isto qualitativamente). Fica por conta do leitor demonstrar que a hamiltoniana (7.2.4) é a "energia total" no sistema de referência girante. ■

Também pode acontecer o oposto, isto é, H coincidir com a energia total mas não se conservar devido a uma dependência temporal explícita (Goldstein 1980).

Há uma importante distinção de comportamento entre a lagrangiana e a hamiltoniana no que tange à escolha de coordenadas. Diferentes escolhas de coordenadas generalizadas alteram a *forma funcional* de uma lagrangiana, mas o seu valor permanece o mesmo. No que se refere à hamiltoniana, todavia, tanto a *forma funcional* quanto o *valor* de H dependem do conjunto de coordenadas generalizadas adotado. Assim, pode muito bem acontecer de a quantidade representada por H ser conservada num sistema de coordenadas e, num outro, variar com o tempo. Esta mutabilidade de H, que à primeira vista pode parecer inconveniente, é altamente vantajosa e constitui a base do formalismo de Hamilton-Jacobi, um poderoso método alternativo de integração das equações de movimento que será estudado num dos próximos capítulos.

7.3 Teorema do Virial

Seja f uma função real de uma variável real t. O valor médio de f é definido por

$$\langle f \rangle = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau f(t)\, dt \tag{7.3.1}$$

sempre que este limite existe. Em particular, quando f é uma função periódica e integrável num período o limite existe e é dado simplesmente por

$$\frac{1}{\tau_0} \int_0^{\tau_0} f(t)\, dt \ ,$$

onde, na expressão acima, τ_0 denota o período de f.

Considere, agora, um sistema mecânico com n graus de liberdade descrito pelas variáveis canônicas (q, p) e pela hamiltoniana $H(q, p, t)$. A quantidade

$$\mathcal{V} = -\sum_i q_i \frac{\partial H}{\partial q_i} \tag{7.3.2}$$

é uma versão generalizada do *virial*, introduzido na teoria cinética dos gases na forma $\mathcal{V} = \sum_k \mathbf{r}_k \cdot \mathbf{F}_k$ por Clausius.[3] Em certas circunstâncias, o valor médio de $\mathcal{V}$ pode ser relacionado ao valor médio de outras grandezas dinâmicas do sistema.

Teorema do Virial. Se $q_i(t)$ e $p_i(t)$ são funções limitadas[4] do tempo e se os valores médios de $\sum_i q_i\, \partial H/\partial q_i$ e $\sum_i p_i\, \partial H/\partial p_i$ existem separadamente, então eles são iguais:

$$\left\langle \sum_i q_i \frac{\partial H}{\partial q_i} \right\rangle = \left\langle \sum_i p_i \frac{\partial H}{\partial p_i} \right\rangle \ . \tag{7.3.3}$$

Demonstração. Defina a função

[3] Alguns autores referem-se à quantidade $-(1/2)(\sum_k \mathbf{r}_k \cdot \mathbf{F}_k)$ como o *virial de Clausius*.

[4] Uma função f é dita limitada se existe um número real positivo M tal que $|f(t)| \leq M$ para qualquer t pertencente ao domínio de f.

$$G(t) = \sum_i p_i(t)\, q_i(t) \ . \tag{7.3.4}$$

Como q_i e p_i são funções limitadas do tempo, G também o é. Diferenciando a Eq.(7.3.4) em relação ao tempo obtemos

$$\frac{dG}{dt} = \sum_i p_i\, \dot{q}_i + \sum_i \dot{p}_i\, q_i = \sum_i p_i \frac{\partial H}{\partial p_i} - \sum_i q_i \frac{\partial H}{\partial q_i} \ , \tag{7.3.5}$$

onde usamos as equações de Hamilton. Tomando o valor médio em ambos os lados desta equação resulta

$$\left\langle \frac{dG}{dt} \right\rangle = \left\langle \sum_i p_i \frac{\partial H}{\partial p_i} - \sum_i q_i \frac{\partial H}{\partial q_i} \right\rangle \ . \tag{7.3.6}$$

Entretanto

$$\left\langle \frac{dG}{dt} \right\rangle = \lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau \frac{dG}{dt}\, dt = \lim_{\tau\to\infty} \frac{G(\tau) - G(0)}{\tau} = 0 \tag{7.3.7}$$

porque o numerador é limitado mas o denominador cresce indefinidamente. Com as demais hipóteses do enunciado do teorema[5] somos conduzidos à Eq.(7.3.3), conforme queríamos demonstrar. □

Partícula num Potencial Central

Como primeira aplicação, considere uma partícula executando um movimento espacialmente limitado sob a ação de uma força central, tal como um planeta movendo-se em volta do Sol ou um elétron orbitando um próton no modelo de Rutherford do átomo de hidrogênio. A hamiltoniana em coordenadas cartesianas é

$$H = T + V = \frac{\mathbf{p}^2}{2m} + V(r) \tag{7.3.7}$$

e o teorema do virial implica

$$\left\langle \mathbf{p} \cdot \frac{\mathbf{p}}{m} \right\rangle = \langle \mathbf{r} \cdot \boldsymbol{\nabla} V(r) \rangle = \left\langle \mathbf{r} \cdot \left(\frac{dV}{dr} \frac{\mathbf{r}}{r} \right) \right\rangle = \left\langle r \frac{dV}{dr} \right\rangle \ . \tag{7.3.8}$$

Para o potencial

$$V(r) = \frac{A}{r^n} \tag{7.3.9}$$

a Eq.(7.3.8) torna-se

$$\langle T \rangle = -\frac{n}{2} \langle V \rangle \ . \tag{7.3.10}$$

[5]Vale notar que $\langle f + g \rangle = \langle f \rangle + \langle g \rangle$ só é verdade quando os valores médios de f e g existem separadamente. Examine o seguinte exemplo: $f(t) = 1 + t$ e $g(t) = 1 - t$.

No caso de um potencial gravitacional ou eletrostático ($n = 1$) o teorema do virial estabelece que

$$2\langle T\rangle = -\langle V\rangle \ , \tag{7.3.11}$$

ou seja, o valor médio da energia potencial é menos o dobro do valor médio da energia cinética. Em geral, a energia total é dada por

$$E = \langle E\rangle = \langle T\rangle + \langle V\rangle = \frac{n-2}{n}\langle T\rangle \ . \tag{7.3.12}$$

Para $n = 1$ (caso gravitacional ou eletrostático), esta última equação dá $E = -\langle T\rangle$, mostrando que os estados ligados (órbitas elípticas) têm energia total negativa. Se $n > 2$, no entanto, os possíveis estados ligados do potencial $V(r)$ têm energia total positiva. O teorema do virial tem outras aplicações interessantes, como, por exemplo, a obtenção de soluções aproximadas para as freqüências de osciladores anarmônicos (Sivardière 1986).

Equação de Estado de um Gás Ideal

Aliado ao teorema da equipartição da energia, o teorema do virial propicia uma dedução surpreendentemente concisa da equação de estado de um gás ideal. Consideremos o gás como constituído por um vasto número de partículas sem interação mútua confinadas num recipiente de volume V. As moléculas do gás só experimentam forças quando colidem com as paredes do recipiente, e a força média sobre aquelas que atingem o elemento de área dS é $d\mathbf{F} = -P\hat{\mathbf{n}}dS$, onde P é a pressão do gás e $\hat{\mathbf{n}}$ é o vetor unitário normal *exterior* à superfície S que limita o volume V. Em coordenadas cartesianas a hamiltoniana é $H = \sum_i \mathbf{p}_i^2/2m_i + V(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ onde V faz o papel de potencial das forças exercidas pelas paredes do recipiente. O teorema do virial (7.3.3) aplicado ao gás toma a forma

$$\left\langle \sum_i \mathbf{p}_i \cdot \frac{\mathbf{p}_i}{m_i} \right\rangle = -\left\langle \sum_i \mathbf{r}_i \cdot \mathbf{F}_i \right\rangle = -\oint_S \mathbf{r} \cdot d\mathbf{F} = \oint_S P\mathbf{r} \cdot \hat{\mathbf{n}}\, dS \ , \tag{7.3.13}$$

onde usamos $\partial H/\partial \mathbf{r}_i = \partial V/\partial \mathbf{r}_i = -\mathbf{F}_i$ e levamos em conta que a forças sobre as moléculas do gás devem-se às colisões com as paredes do recipiente. Como P é constante, o teorema da divergência fornece

$$2\,\langle T\rangle = P\int_V \boldsymbol{\nabla}\cdot\mathbf{r}\, dV = P\int_V 3\, dV = 3PV \ . \tag{7.3.14}$$

Pelo teorema da equipartição da energia, a energia cinética média de cada molécula é $3k\,\Theta/2$, onde k é a constante de Boltzmann e Θ é a temperatura absoluta. Assim, se o número de moléculas do gás é N, a equação (7.3.14) toma a forma

$$PV = N\,k\,\Theta \ , \tag{7.3.15}$$

que é a equação de estado de um gás ideal. Se levarmos em conta as interações entre as moléculas do gás aparece um termo adicional no segundo membro da Eq.(7.3.15), o qual depende do potencial de interação e de uma função $g(r)$ que mede a probabilidade de se encontrar duas partículas separadas por uma distância r (Pathria 1972).

7.4 Formulação Hamiltoniana Relativística

Comecemos com o formalismo não-covariante, em que trabalha-se num referencial específico e o tempo, encarado como um parâmetro que descreve a evolução do sistema, é tratado distintamente das coordenadas espaciais. Para uma partícula num potencial V independente das velocidades,

$$L = -mc^2\sqrt{1 - v^2/c^2} - V(\mathbf{r}) \ , \tag{7.4.1}$$

donde

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = \frac{m\,\mathbf{v}}{\sqrt{1 - v^2/c^2}} \ . \tag{7.4.2}$$

Note que o momento canônico $\mathbf{p}$ coincide com o momento linear relativístico. Portanto,

$$H = \mathbf{v}\cdot\mathbf{p} - L = \frac{m\,v^2}{\sqrt{1 - v^2/c^2}} + mc^2\sqrt{1 - v^2/c^2} + V(\mathbf{r}) = \frac{m\,c^2}{\sqrt{1 - v^2/c^2}} + V \ , \tag{7.4.3}$$

e H é a energia total, embora a lagrangiana (7.4.1) não tenha a forma de diferença entre energia cinética e energia potencial. É preciso, no entanto, exprimir H como função de $(\mathbf{r}, \mathbf{p})$. Isto é facilmente conseguido recorrendo à relação relativística (6.6.38) entre energia e momento linear de uma partícula livre, da qual resulta

$$H = \sqrt{p^2c^2 + m^2c^4} + V \ . \tag{7.4.4}$$

No caso de uma partícula num campo eletromagnético externo,

$$L = -mc^2\sqrt{1 - v^2/c^2} - e\,\phi(\mathbf{r}, t) + \frac{e}{c}\,\mathbf{v}\cdot\mathbf{A}(\mathbf{r}, t) \ , \tag{7.4.5}$$

donde

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = \frac{m\,\mathbf{v}}{\sqrt{1 - v^2/c^2}} + \frac{e}{c}\,\mathbf{A} \ . \tag{7.4.6}$$

Note que, agora, o momento canônico não é mais idêntico ao momento linear relativístico. Temos

$$H = \mathbf{v}\cdot\mathbf{p} - L = \frac{m\,c^2}{\sqrt{1 - v^2/c^2}} + e\,\phi \ , \tag{7.4.7}$$

e novamente H é a energia total no caso de campos $\mathbf{E}$ e $\mathbf{B}$ estáticos. Recorrendo outra vez a (6.6.38) com o momento linear relativístico substituído por $\mathbf{p} - (e/c)\mathbf{A}$, obtemos a hamiltoniana em sua forma final:

$$H(\mathbf{r}, \mathbf{p}, t) = \sqrt{\left(\mathbf{p} - \frac{e}{c}\,\mathbf{A}(\mathbf{r}, t)\right)^2 c^2 + m^2c^4} + e\,\phi(\mathbf{r}, t) \ . \tag{7.4.8}$$

A hamiltoniana relativística de uma partícula livre é dada por (7.4.4) com $V = 0$, e claramente a transição para a hamiltoniana (7.4.8) de uma partícula num campo eletromagnético externo se faz pela prescrição

$$H \longrightarrow H - e\phi \quad , \quad \mathrm{p} \longrightarrow \mathrm{p} - \frac{e}{c}\mathbf{A} \ , \tag{7.4.9}$$

conhecida como *acoplamento mínimo*.

Quanto à formulação covariante, tomando como ponto de partida a lagrangiana

$$L = \frac{m}{2}U^\alpha U_\alpha + \frac{e}{c}A^\alpha U_\alpha \tag{7.4.10}$$

para uma partícula num campo eletromagnético externo, obtemos os momentos canônicos

$$P^\mu = \frac{\partial L}{\partial U_\mu} = mU^\mu + \frac{e}{c}A^\mu \ . \tag{7.4.11}$$

O quadrimomento canônico não coincide com o quadrimomento da partícula. A hamiltoniana covariante é dada por

$$H = P^\alpha U_\alpha - L = \frac{1}{2m}\Big(P^\alpha - \frac{e}{c}A^\alpha\Big)\Big(P_\alpha - \frac{e}{c}A_\alpha\Big) \ . \tag{7.4.12}$$

A passagem da hamiltoniana de uma partícula livre para a de uma partícula num campo eletromagnético externo se realiza pelo acoplamento mínimo

$$P^\alpha \longrightarrow P^\alpha - \frac{e}{c}A^\alpha \ , \tag{7.4.13}$$

que é a versão covariante de (7.4.9). As equações de Hamilton covariantes são

$$\frac{dx^\mu}{d\tau} = \frac{\partial H}{\partial P_\mu} \quad , \quad \frac{dP^\mu}{d\tau} = -\frac{\partial H}{\partial x_\mu} \ . \tag{7.4.14}$$

■ **Exercício 7.4.1.** (a) Prove que a parte espacial das Eqs.(7.4.14) equivale à equação de movimento relativística $d(\gamma m\mathbf{v})/dt = e(\mathbf{E} + \mathbf{v}\times\mathbf{B}/c)$. (b) A hamiltoniana (7.4.12) é um escalar, logo não pode ser igual à energia total da partícula. Mostre que, apesar disso, H é uma constante de movimento e seu valor é $mc^2/2$. ■

No que diz respeito às componentes temporais de (7.4.14), temos

$$U^0 = \frac{\partial H}{\partial P_0} = \frac{1}{m}\Big(P^0 - \frac{e}{c}A^0\Big) \ , \tag{7.4.15}$$

donde

$$P^0 = mU^0 + \frac{e}{c}A^0 = \frac{E + e\phi}{c} \ . \tag{7.4.16}$$

De acordo com esta equação, a energia total $E_{tot} = E + e\phi$ desempenha o papel de momento conjugado ao tempo.[6] Temos, ainda,

$$\frac{dP_0}{d\tau} = -\frac{\partial H}{\partial x^0} = -\frac{1}{c}\frac{\partial H}{\partial t} = \frac{e}{mc^2}\Big(p^\alpha - \frac{e}{c}A^\alpha\Big)\frac{\partial A_\alpha}{\partial t} \ , \tag{7.4.17}$$

ou, equivalentemente,

$$\gamma\frac{d}{dt}(E + e\,\phi) = \frac{e}{mc}\,mU^\alpha\frac{\partial A_\alpha}{\partial t} = \frac{e}{c}\left[\gamma c\frac{\partial\phi}{\partial t} - \gamma\mathbf{v}\cdot\frac{\partial\mathbf{A}}{\partial t}\right] \ . \tag{7.4.18}$$

Cancelando um fator γ comum, esta última equação pode ser reescrita na forma

$$\frac{dE}{dt} + e\left(\mathbf{v}\cdot\boldsymbol{\nabla}\phi + \frac{\partial\phi}{\partial t}\right) = e\,\frac{\partial\phi}{\partial t} - \frac{e\mathbf{v}}{c}\cdot\frac{\partial\mathbf{A}}{\partial t} \ , \tag{7.4.19}$$

ou, finalmente,

$$\frac{dE}{dt} = e\left(-\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{\partial\mathbf{A}}{\partial t}\right)\cdot\mathbf{v} = e\,\mathbf{E}\cdot\mathbf{v} = \mathbf{F}\cdot\mathbf{v} \ , \tag{7.4.20}$$

onde $\mathbf{F} = e(\mathbf{E} + \mathbf{v}\times\mathbf{B}/c)$. A Eq.(7.4.20) exprime corretamente a lei da conservação da energia, completando a verificação de que as equações de Hamilton covariantes (7.4.14) têm o mesmo conteúdo que a equação de movimento covariante (6.6.24).

7.5 Forma Variacional das Equações de Hamilton

O princípio de Hamilton, formulado em termos de variações da trajetória no espaço de configuração, escreve-se

$$\delta S \equiv \delta\int_{t_1}^{t_2} L(q,\dot{q},t)\,dt = 0 \ , \tag{7.5.1}$$

com $\delta q_i(t_1) = \delta q_i(t_2) = 0$. Este princípio variacional pode ser transcrito em termos de variações da trajetória no espaço de fase da seguinte forma:

$$\delta S \equiv \delta\int_{t_1}^{t_2}\Big\{\sum_{i=1}^{n} p_i\,\dot{q}_i - H(q,p,t)\Big\}\,dt = 0 \ , \tag{7.5.2}$$

onde $\delta q_i(t_1) = \delta q_i(t_2) = 0$ e as variações dos qs e ps são consideradas independentes. Com efeito, temos

$$\delta S = \int_{t_1}^{t_2} dt\,\sum_{i=1}^{n}\Big(p_i\,\delta\dot{q}_i + \dot{q}_i\,\delta p_i - \frac{\partial H}{\partial q_i}\,\delta q_i - \frac{\partial H}{\partial p_i}\,\delta p_i\Big) = 0 \ . \tag{7.5.3}$$

Uma integração por partes fornece

[6] Num certo sentido, este resultado não é uma peculiaridade da mecânica relativística, conforme se discutirá na Seção 7.6.

$$\int_{t_1}^{t_2} p_i\,\delta\dot{q}_i\,dt = p_i\,\delta q_i\Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} \dot{p}_i\,\delta q_i\,dt = -\int_{t_1}^{t_2} \dot{p}_i\,\delta q_i\,dt \ , \tag{7.5.4}$$

de modo que a Eq.(7.5.3) torna-se

$$\delta S = \int_{t_1}^{t_2} dt \sum_{i=1}^{n} \left\{ \left(\dot{q}_i - \frac{\partial H}{\partial p_i}\right)\delta p_i - \left(\dot{p}_i + \frac{\partial H}{\partial q_i}\right)\delta q_i \right\} = 0 \ . \tag{7.5.5}$$

Uma vez que as variações dos qs e ps são arbitrárias e independentes entre si, decorre de (7.5.5) que

$$\dot{q}_i - \frac{\partial H}{\partial p_i} = 0 \quad , \quad \dot{p}_i + \frac{\partial H}{\partial q_i} = 0 \quad , \quad i = 1, \ldots, n \ , \tag{7.5.6}$$

que são as equações de Hamilton.

■ **Exercício 7.5.1.** Encarando a ação S como

$$S[q,p] = \int_{t_1}^{t_2} F(q,p,\dot{q},\dot{p},t)\,dt$$

com $F = \sum_i p_i\,\dot{q}_i - H(q,p,t)$ e aplicando as equações de Euler à função F, reobtenha as equações de Hamilton a partir da condição de que S seja mínima (mais geralmente, estacionária) para a trajetória física. ■

O princípio variacional (7.5.2) requer apenas que as variações das coordenadas anulem-se nas extremidades do intervalo de integração, as variações dos momentos permanecendo inteiramente arbitrárias. Há diversas vantagens em exigir que também as variações dos momentos anulem-se nas extremidades do intervalo de integração. Por exemplo, uma integração por partes em (7.5.2) conduz a

$$\delta S = \delta\left\{ \int_{t_1}^{t_2} dt\Big[-\sum_{i=1}^{n} q_i\,\dot{p}_i - H(q,p,t)\Big] + \sum_{i=1}^{n} q_i\,p_i\Big|_{t_1}^{t_2} \right\} = 0 \ . \tag{7.5.7}$$

Mas $\delta(q_i p_i) = 0$ em t_1 e t_2 se os δqs e δps anulam-se em t_1 e t_2, de forma que o novo princípio variacional

$$\delta \int_{t_1}^{t_2} dt\Big[-\sum_{i=1}^{n} q_i\,\dot{p}_i - H(q,p,t)\Big] = 0 \tag{7.5.8}$$

também fornece as equações de Hamilton. Combinando (7.5.2) e (7.5.8) obtemos um terceiro princípio variacional

$$\delta \int_{t_1}^{t_2} dt \left\{ \frac{1}{2}\sum_{i=1}^{n} (p_i\,\dot{q}_i - q_i\,\dot{p}_i) - H(q,p,t) \right\} = 0 \tag{7.5.9}$$

que gera as equações de Hamilton, com a virtude de oferecer um tratamento mais simétrico às coordenadas e momentos. Daqui por diante suporemos que em todos os princípios variacionais formulados no espaço de fase as variações das coordenadas e momentos são nulas nos extremos temporais, o que revelar-se-á particularmente conveniente no estudo das transformações canônicas, a ser empreendido no próximo capítulo.

7.6 O Tempo Como Variável Canônica

Um fenômeno interessante ocorre quando tentamos incluir o tempo no rol das coordenadas generalizadas com o emprego de uma parâmetro auxiliar θ capaz de descrever a evolução do sistema. No caso de uma partícula, θ poderia ser a distância percorrida ao longo de sua trajetória no espaço. A única restrição sobre θ é $d\theta/dt > 0$, isto é, t deve ser uma função crescente de θ. Façamos, portanto, $q_{n+1} \equiv t$ e escrevamos a ação na forma

$$S = \int_{\theta_1}^{\theta_2} \Big\{ \sum_{i=1}^{n} p_i \frac{dq_i}{dt} - H(q,p,t) \Big\} \frac{dt}{d\theta}\, d\theta = \int_{\theta_1}^{\theta_2} \Big\{ \sum_{i=1}^{n} p_i \frac{dq_i}{d\theta} - H \frac{dq_{n+1}}{d\theta} \Big\} d\theta \qquad (7.6.1)$$

ou, equivalentemente,

$$S = \int_{\theta_1}^{\theta_2} d\theta \sum_{i=1}^{n+1} p_i\, q_i' \ , \qquad (7.6.2)$$

onde

$$q_i' \equiv \frac{dq_i}{d\theta} \quad , \quad i = 1, \ldots, n+1 \quad , \quad p_{n+1} = -H \ . \qquad (7.6.3)$$

Três características das Eqs.(7.6.2) e (7.6.3) merecem destaque: (i) o momento conjugado ao tempo é $-H$; (ii) no *espaço de fase estendido* $(q,p) \equiv (q_1, \ldots, q_{n+1}, p_1, \ldots, p_{n+1})$ a hamiltoniana é zero; (iii) a equação $p_{n+1} = -H$ é um vínculo no espaço de fase estendido.

Introduzindo à função super-hamiltoniana[7]

$$\mathcal{H}(q,p) = p_{n+1} + H(q_1, \ldots, q_{n+1}, p_1, \ldots, p_n) \ , \qquad (7.6.4)$$

definida no espaço de fase estendido, o vínculo exprime-se como

$$\mathcal{H}(q,p) = 0 \ . \qquad (7.6.5)$$

Por causa deste vínculo, para se formular corretamente as equações de movimento pelo método variacional no espaço de fase estendido é preciso fazer uso de um multiplicador de Lagrange λ e reescrever a ação na forma (vide o último parágrafo da Seção 2.4)

$$S[q,p,\lambda] = \int_{\theta_1}^{\theta_2} d\theta \left\{ \sum_{i=1}^{n+1} p_i\, q_i' - \lambda \mathcal{H}(q,p) \right\} \equiv \int_{\theta_1}^{\theta_2} d\theta\, F(q,p,\lambda,q',p',\lambda') \ . \qquad (7.6.6)$$

As equações de Euler decorrentes de $\delta S = 0$ são

$$\frac{d}{d\theta}\Big(\frac{\partial F}{\partial q_k'}\Big) - \frac{\partial F}{\partial q_k} = 0 \quad , \quad \frac{d}{d\theta}\Big(\frac{\partial F}{\partial p_k'}\Big) - \frac{\partial F}{\partial p_k} = 0 \quad , \quad k = 1, \ldots, n+1 \ , \qquad (7.6.7)$$

[7] Este nome é tomado de empréstimo da cosmologia quântica.

e, também,

$$\frac{d}{d\theta}\left(\frac{\partial F}{\partial \lambda'}\right) - \frac{\partial F}{\partial \lambda} = 0 \ . \tag{7.6.8}$$

Levando em conta a forma explícita de F, estas equações reduzem-se a

$$p'_k + \lambda \frac{\partial \mathcal{H}}{\partial q_k} = 0 \quad , \quad -q'_k + \lambda \frac{\partial \mathcal{H}}{\partial p_k} = 0 \quad , \quad k = 1, \ldots, n+1 \ , \tag{7.6.9}$$

$$\mathcal{H}(q,p) = 0 \ . \tag{7.6.10}$$

Portanto, o vínculo é automaticamente satisfeito em conseqüência do princípio variacional. Considerando a $(n+1)$-ésima componente das Eqs.(7.6.9), deduzimos

$$-\frac{dt}{d\theta} + \lambda = 0 \quad \Longrightarrow \quad \lambda = \frac{dt}{d\theta} \ ; \tag{7.6.11}$$

além disso,

$$\frac{dp_{n+1}}{d\theta} + \lambda \frac{\partial \mathcal{H}}{\partial q_{n+1}} = 0 \ \Longrightarrow \ -\frac{dH}{d\theta} = -\frac{dt}{d\theta}\frac{\partial H}{\partial t} \ \Longrightarrow \ \frac{dH}{dt} = \frac{\partial H}{\partial t} \ , \tag{7.6.12}$$

que não é outra coisa senão a Eq.(7.2.2). Por fim, com o resultado (7.6.11) as n primeiras componentes de (7.6.9) tornam-se

$$\frac{dq_k}{d\theta} = \lambda \frac{\partial H}{\partial p_k} \ \Longrightarrow \ \frac{dq_k}{dt}\frac{dt}{d\theta} = \frac{dt}{d\theta}\frac{\partial H}{\partial p_k} \ \Longrightarrow \ \dot{q}_k = \frac{\partial H}{\partial p_k} \ , \tag{7.6.13}$$

$$\frac{dp_k}{d\theta} = -\lambda \frac{\partial H}{\partial q_k} \ \Longrightarrow \ \frac{dp_k}{dt}\frac{dt}{d\theta} = -\frac{dt}{d\theta}\frac{\partial H}{\partial q_k} \ \Longrightarrow \ \dot{p}_k = -\frac{\partial H}{\partial q_k} \ , \tag{7.6.14}$$

e recuperamos as equações de Hamilton.

Em suma, pode-se tratar o tempo em igualdade de condições com as coordenadas generalizadas desde que se esteja disposto a lidar com um espaço de fase estendido no qual há um vínculo, isto é, as variáveis canônicas não são mutuamente independentes. Um subproduto formal desse tratamento é que o tempo e a hamiltoniana aparecem como variáveis canonicamente conjugadas. Este resultado tem repercussões que transcendem a mecânica clássica, pois sugere a existência de uma relação de incerteza entre energia e tempo na mecânica quântica, o que de fato se verifica.

■ **Exercício 7.6.1.** Podemos interpretar a Eq.(7.6.11) não como determinando λ, mas como servindo para determinar θ. Em outras palavras, uma escolha de λ define que parâmetro θ está sendo usado para descrever o desenvolvimento da trajetória no espaço de fase. Por exemplo, a escolha $\lambda = 1$ equivale a tomar $\theta = t$, e o tempo usual é o parâmetro que está sendo empregado para descrever a evolução do sistema. Mostre que, neste caso, as Eqs.(7.6.9) e (7.6.10) coincidem exatamente com as equações de Hamilton (7.1.9) acrescidas de (7.2.2). ■

Sistemas com o Tempo Parametrizado

Para que a hamiltoniana de um sistema seja zero, a condição necessária e suficiente é que a lagrangiana seja uma função homogênea de primeiro grau das velocidades. Isto é uma conseqüência imediata de (7.1.2), (7.1.5) e do teorema de Euler das funções homogêneas. Define-se um sistema com o tempo parametrizado como aquele cujas equações de movimento independem do parâmetro de evolução escolhido ou, equivalentemente, cuja ação é invariante sob transformações arbitrárias do parâmetro de evolução temporal. Um sistema com o tempo parametrizado é necessariamente descrito por uma lagrangiana que é função homogênea de primeiro grau das velocidades generalizadas.

■ **Exercício 7.6.2.** Um sistema com o tempo parametrizado caracteriza-se por uma ação invariante sob reparametrizações do tempo, isto é, mudanças arbitrárias do parâmetro de evolução temporal. Considere, portanto, um sistema com lagrangiana $L(q,\dot{q})$ sem dependência temporal explícita e suponha que a ação correspondente seja invariante sob uma mudança infinitesimal arbitrária $t' = t + \epsilon X(t)$ do parâmetro de evolução. Usando a condição de Noether (2.7.8), prove que L é função homogênea de primeiro grau das velocidades. ■

Em problemas invariantes sob transformações do parâmetro de evolução temporal, o procedimento para se passar do espaço de fase estendido, onde há o vínculo super-hamiltoniano $\mathcal{H}(q,p) = 0$, para o *espaço de fase reduzido*, no qual as variáveis canônicas são mutuamente independentes, consiste em percorrer o caminho inverso do que conduziu a (7.6.2). Dada uma ação da forma (7.6.6), *escolhe-se* uma variável canônica para desempenhar o papel do tempo e resolve-se a equação $\mathcal{H}(q,p) = 0$ para a variável canonicamente conjugada correspondente. Esta variável conjugada ao tempo, agora expressa em função das demais variáveis canônicas, deve ser introduzida na ação. A coordenada escolhida como tempo e o seu momento conjugado não mais aparecem na ação reduzida resultante, que tem a forma

$$S = \int_{t_1}^{t_2} \Big\{ \sum_{i=1}^{n} p_i \dot{q}_i - H(q,p,t) \Big\} dt \ , \tag{7.6.15}$$

a qual permite identificar a hamiltoniana não-nula $H(q,p,t)$ no espaço de fase reduzido.

À guisa de ilustração desse procedimento, consideremos a partícula livre relativística. Usando a lagrangiana (6.8.4) com $V = 0$ podemos escrever a ação na forma

$$S = -mc^2 \int dt \sqrt{1 - v^2/c^2} = -mc \int \sqrt{c^2\, dt^2 - dx^2 - dy^2 - dz^2} \tag{7.6.16}$$

ou, em termos de um parâmetro escalar arbitrário θ,

$$S = -mc \int_{\theta_1}^{\theta_2} \sqrt{V^\alpha V_\alpha}\, d\theta \tag{7.6.17}$$

onde

$$V^\alpha = \frac{dx^\alpha}{d\theta} \ . \tag{7.6.18}$$

Note que a lagrangiana covariante

$$L = -mc\sqrt{V^\alpha V_\alpha} \tag{7.6.19}$$

é homogênea de primeiro grau em V^α. O quadrimomento canônico é

$$P^\mu = -\frac{\partial L}{\partial V_\mu} = mc\frac{V^\mu}{\sqrt{V^\alpha V_\alpha}} \tag{7.6.20}$$

e satisfaz a equação

$$P^\mu P_\mu = m^2c^2 \ , \tag{7.6.21}$$

que constituiu um vínculo no espaço de fase estendido (o sinal negativo na definição do quadrimomento canônico visa torná-lo idêntico ao quadrimomento usual quando $\theta = \tau$). Temos a hamiltoniana[8]

$$H = -P^\mu V_\mu - L = -mc\frac{V^\mu V_\mu}{\sqrt{V^\alpha V_\alpha}} + mc\sqrt{V^\alpha V_\alpha} = 0 \ , \tag{7.6.22}$$

conforme havíamos antecipado. A ação adequada no espaço de fase estendido escreve-se

$$S = \int_{\theta_1}^{\theta_2} d\theta \left[-P^\mu V_\mu + \lambda(P^2 - m^2c^2)\right] \ , \tag{7.6.23}$$

que tem a forma (7.6.6) com $\mathcal{H} = P^2 - m^2c^2$. Variando independentemente x^α, P^α e λ, esta ação dá origem às equações de movimento

$$\frac{dx^\alpha}{d\theta} = 2\lambda P^\alpha \quad , \quad \frac{dP^\alpha}{d\theta} = 0 \ , \tag{7.6.24}$$

suplementadas por (7.6.21). De (7.6.24) e (7.6.21) se deduz facilmente

$$\lambda = \frac{1}{2mc}\sqrt{V^\alpha V_\alpha} \ , \tag{7.6.25}$$

da qual vê-se, por exemplo, que a escolha $\lambda = 1/2m$ corresponde a tomar θ como o tempo próprio τ, pois neste caso V^μ satisfaz (6.6.7), de modo que $V^\mu = U^\mu$ onde U^μ é a quadrivelocidade. Por outro lado, as Eqs.(7.6.24) equivalem a

$$\frac{d}{d\theta}\left(\frac{V^\alpha}{2\lambda}\right) = 0 \ , \tag{7.6.26}$$

ou, recorrendo a (7.6.25),

$$\frac{d}{d\theta}\left(\frac{mcV^\alpha}{\sqrt{V^\beta V_\beta}}\right) = 0 \ , \tag{7.6.27}$$

[8] O sinal negativo no termo $P^\mu V_\mu$ em (7.6.22) e (7.6.23) é para compensar o sinal negativo introduzido na definição (7.6.20) do quadrimomento canônico. Em conseqüência, os sinais se invertem nas equações de Hamilton: $V^\alpha = -\partial H/\partial P_\alpha$, $\dot{P}^\alpha = \partial H/\partial x_\alpha$.

que é exatamente a equação de movimento gerada pela lagrangiana (7.6.19). Em particular, escolhendo-se θ igual ao tempo próprio τ esta equação reduz-se a $dP^\mu/d\tau = 0$ com $P^\mu = mU^\mu$, que é o esperado para uma partícula livre.

A passagem ao espaço de fase reduzido pode ser realizada escolhendo $\theta = x^0/c \equiv t$ e resolvendo a equação de vínculo (7.6.21) para a variável canonicamente conjugada cP_0:

$$(P^0)^2 - \mathbf{p}^2 = m^2c^2 \implies cP^0 = cP_0 = \sqrt{\mathbf{p}^2c^2 + m^2c^4} \ . \tag{7.6.28}$$

A solução negativa para P^0 deve ser descartada porque a energia $E = cP^0$ de uma partícula livre é positiva. Levando este resultado na ação (7.6.23) resulta

$$S = \int_{t_1}^{t_2} dt \left\{ -P^0 \frac{dx_0}{dt} + \mathbf{p} \cdot \frac{d\mathbf{r}}{dt} \right\} = \int_{t_1}^{t_2} dt \left\{ \mathbf{p} \cdot \dot{\mathbf{r}} - \sqrt{\mathbf{p}^2c^2 + m^2c^4} \right\} \ , \tag{7.6.29}$$

que é da forma (7.6.15) com a hamiltoniana correta

$$H = \sqrt{\mathbf{p}^2c^2 + m^2c^4} \ , \tag{7.6.30}$$

que não é outra coisa senão a energia (6.6.38) de uma partícula livre relativística.

O procedimento de redução descrito acima é a base do formalismo de Arnowitt-Deser-Misner (ADM), desenvolvido com vistas à formulação hamiltoniana da teoria geral da relatividade (Arnowitt, Deser & Misner 1962). A teoria da gravitação de Einstein é invariante sob transformações gerais de coordenadas e não há nenhuma variável que se preste naturalmente ao papel de tempo. Trata-se, portanto, de uma teoria na qual o tempo é parametrizado, a hamiltoniana é zero e a ação é da forma (7.6.6). O formalismo ADM tem se revelado importante nas tentativas de construção de um teoria quântica da gravitação, e tem encontrado muitas aplicações na cosmologia quântica.

■ **Exemplo 7.6.1.** (Lemos 1996) Num problema de cosmologia quântica a ação tem a forma

$$S = \int_{t_1}^{t_2} dt \left\{ \dot{R} p_R + \dot{\phi} p_\phi - N\Big(\frac{p_R^2}{24R} - \frac{p_\phi^2}{2R^3} + 6kR \Big) \right\} \ , \tag{7.6.31}$$

onde k é uma constante e t designa o parâmetro que descreve a evolução do sistema. Sabendo que N e R são quantidades estritamente positivas, obter as equações de movimento no espaço de fase estendido (R, ϕ, p_R, p_ϕ) e a hamiltoniana no espaço de fase reduzido resultante da escolha $t = \phi$.

Solução. A ação (7.6.31) tem a forma (7.6.6) com N desempenhando o papel de um multiplicador de Lagrange. A variável R é interpretada como o raio do universo, e um R crescente corresponde a um universo em expansão. Nesse modelo simplificado, $\phi(t)$ é um campo escalar homogêneo que representa o conteúdo material do universo e é responsável por sua expansão. Impondo $\delta S = 0$ com as variações de R, ϕ, p_R, p_ϕ, N arbitrárias e independentes, resultam as equações de movimento no espaço de fase estendido

$$\dot{R} = \frac{N p_R}{12R} \ , \quad \dot{p}_R = N\Big(\frac{p_R^2}{24R^2} - \frac{3p_\phi^2}{2R^4} - 6k \Big) \ , \quad \dot{\phi} = -\frac{N p_\phi}{R^3} \ , \quad \dot{p}_\phi = 0 \ , \tag{7.6.32}$$

suplementadas pelo vínculo super-hamiltoniano

$$\frac{p_R^2}{24R} - \frac{p_\phi^2}{2R^3} + 6kR = 0 \ . \tag{7.6.33}$$

Na relatividade geral o parâmetro de evolução t é arbitrário e, como no caso da partícula livre relativística, a fixação do multiplicador de Lagrange N corresponde a uma escolha do tempo, isto é, de que quantidade física está sendo empregada para acompanhar a evolução do sistema. A escolha $t = \phi$ exige que resolvamos a Eq.(7.6.33) para p_ϕ, e a hamiltoniana no espaço de fase reduzido é $H = -p_\phi$. Como agora $\dot{\phi} = 1$, a terceira das Eqs.(7.6.32) mostra que $p_\phi < 0$, pois N e R são grandezas positivas. Resolvendo (7.6.33) para p_ϕ e tomando a raiz quadrada negativa obtém-se a hamiltoniana reduzida

$$H = -p_\phi = \sqrt{\frac{R^2 p_R^2}{12} + 12kR^4} \ . \tag{7.6.34}$$

A ação no espaço de fase reduzido é obtida substituindo $\dot{\phi} = 1$ e a solução (7.6.34) para p_ϕ em (7.6.31), dando como resultado

$$S = \int_{t_1}^{t_2} dt \left\{ \dot{R} p_R - \sqrt{\frac{R^2 p_R^2}{12} + 12kR^4} \right\} \ . \tag{7.6.35}$$

O espaço de fase reduzido é gerado pelas variáveis canônicas (R, p_R). ■

7.7 Princípio de Maupertuis

Historicamente, o primeiro princípio variacional da mecânica foi proposto por Maupertuis,[9] em 1744, numa forma obscura (Yourgrau & Mandelstam 1968; Dugas 1988) e marcada pela intenção de dar uma base não apenas racional mas também teológica à mecânica: em seu *Essai de Cosmologie* (1750), Maupertuis usou o seu princípio variacional como prova do poder e da sabedoria de Deus. Esse princípio ganhou forma definitiva e matematicamente rigorosa graças aos esforços de Euler e Lagrange. Originalmente ele recebeu o nome de "princípio da mínima ação", mas este título é hoje reservado por quase todos os físicos para o princípio de Hamilton. Empregaremos a terminologia sugerida por Landau & Lifchitz (1966), e a quantidade envolvida no princípio de Maupertuis será designada como "ação reduzida".

A variação Δ que ocorre no princípio de Maupertuis envolve uma variação do tempo, ao passo que a variação δ usual é feita a tempo fixo. É necessário, portanto, descrever com clareza o significado da variação Δ.

[9] Matemático e astrônomo, Pierre-Louis Moreau de Maupertuis foi membro da Academia de Ciências de Paris, introduziu a teoria da gravitação newtoniana na França, liderou uma expedição à Lapônia para medições que comprovaram o achatamento da Terra nos pólos, tornou-se membro da Academia de Ciências de Berlin, que presidiu de 1745 a 1753, e chegou a ser chamado de "Sir Isaac Maupertuis" por Voltaire. Maupertuis foi acusado – injustamente, ao que parece – pelo matemático alemão Samuel König de ter plagiado Leibnitz. Na disputa de prioridade em torno do então chamado "princípio da mínima ação", Maupertuis foi apoiado por ninguém menos do que Leonhard Euler, mas Voltaire, seu antigo admirador, voltou-se contra ele e satirizou o "achatador da Terra" tão impiedosamente que Maupertuis abandonou Berlin em 1753.

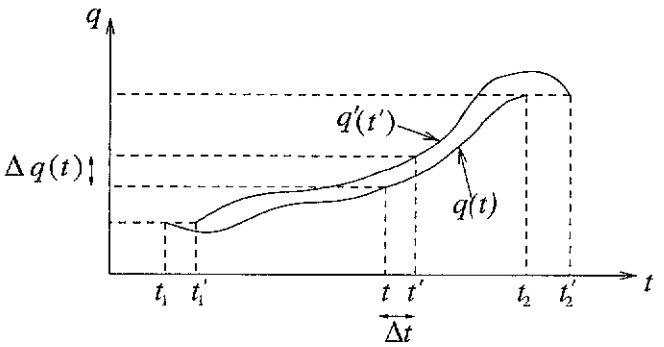


Figura 7.1: Variação envolvida no princípio de Maupertuis.

Definição 7.7.1. A variação Δ é definida por

$$t' = t + \Delta t \equiv t + \epsilon X(t) \ , \tag{7.7.1}$$

$$q_i'(t') = q_i(t) + \Delta q_i(t) \equiv q_i(t) + \epsilon \eta_i(t) \ , \tag{7.7.2}$$

ou seja,

$$\Delta t = t' - t = \epsilon X(t) \quad , \quad \Delta q_i(t) = q_i'(t') - q_i(t) = \epsilon\, \eta_i(t) \ , \tag{7.7.3}$$

onde ϵ é um parâmetro infinitesimal (ver Fig. 7.1).

Princípio de Maupertuis. Suponha que: (i) H não dependa explicitamente do tempo; (ii) H seja conservada e tenha o mesmo valor em todas as trajetórias de comparação; (iii) as configurações inicial e final sejam fixas, isto é, $\Delta q_i(t_1) = \Delta q_i(t_2) = 0$. Então

$$\Delta I \equiv \Delta \int_{t_1}^{t_2} dt \sum_i p_i\, \dot{q}_i = 0 \tag{7.7.4}$$

para uma variação em torno da trajetória física $q(t)$. Em outras palavras, dentre todas as trajetórias imagináveis capazes de conectar duas configurações fixas, com hamiltoniana constante, o movimento físico verdadeiro é aquele para o qual a *ação reduzida*

$$I \equiv \int_{t_1}^{t_2} dt \sum_i p_i\, \dot{q}_i \tag{7.7.5}$$

é mínima ou, mais geralmente, estacionária.

Demonstração. Para a trajetória real a ação reduzida é dada por

$$I \equiv \int_{t_1}^{t_2} dt \sum_i p_i\, \dot{q}_i = \int_{t_1}^{t_2} (L + H)\, dt = \int_{t_1}^{t_2} L\, dt + H\,(t_2 - t_1) \ , \tag{7.7.6}$$

pois H permanece constante durante o movimento. Conforme a hipótese (ii),

$$\Delta I = \Delta \int_{t_1}^{t_2} L\,dt + H\,(\Delta t_2 - \Delta t_1)\ , \tag{7.7.7}$$

onde $\Delta t_1 = \epsilon\, X(t_1)$ com definição análoga para Δt_2. A variação Δ é um caso particular das transformações infinitesimais (2.7.1) admissíveis no teorema de Noether. Aproveitando os cálculos da Seção 2.7 que conduziram às equações (2.7.7) a (2.7.9), obtemos

$$\Delta \int_{t_1}^{t_2} L\,dt = \int_{t_1}^{t_2} \frac{d}{dt}\Big\{\sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i}\Delta q_i - h\Delta t\Big\}dt = \Big(\sum_{i=1}^{n} p_i \Delta q_i - h\Delta t\Big)\Big|_{t_1}^{t_2}\ . \tag{7.7.8}$$

Conseqüentemente, como $h = H$ e $\Delta q_i(t_1) = \Delta q_i(t_2) = 0$, a Eq.(7.7.8) reduz-se a

$$\Delta \int_{t_1}^{t_2} L\,dt = -H\,(\Delta t_2 - \Delta t_1)\ , \tag{7.7.9}$$

onde usamos a constância de H. A inserção deste último resultado em (7.7.7) conduz imediatamente à Eq.(7.7.4), completando a demonstração do princípio de Maupertuis. □

Princípio de Maupertuis e Óptica Geométrica

É instrutivo examinar o conteúdo do princípio de Maupertuis em alguns casos particulares. Quando as equações (1.4.2) que definem as coordenadas generalizadas não dependem explicitamente do tempo, a energia cinética é uma função quadrática das velocidades generalizadas:

$$T = \frac{1}{2}\sum_{k,l} M_{kl}(q)\,\dot{q}_k\,\dot{q}_l\ . \tag{7.7.10}$$

Se, além disso, $L = T - V$ com V independente das velocidades,

$$\sum_i \dot{q}_i\, p_i = \sum_i \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} = 2T \tag{7.7.11}$$

em virtude do teorema de Euler das funções homogêneas. Nessas circunstâncias, o princípio de Maupertuis afirma que

$$\Delta \int_{t_1}^{t_2} T\,dt = 0\ . \tag{7.7.12}$$

Por exemplo, considere uma partícula ou corpo rígido na ausência de forças externas. A energia cinética é constante e o princípio de Maupertuis toma a forma especial

$$\Delta(t_2 - t_1) = 0\ . \tag{7.7.13}$$

Em palavras: de todos os percursos possíveis entre duas configurações fixas, o caminho real percorrido pelo sistema é aquele para o qual o tempo de trânsito é mínimo. A semelhança com o princípio de Fermat da óptica geométrica é flagrante.

Princípio de Jacobi

A forma (7.7.10) da energia cinética sugere que o espaço de configuração seja encarado como um espaço dotado de uma métrica, com a distância $d\rho$ entre dois pontos infinitesimalmente próximos dada por

$$d\rho^2 = \sum_{k,l} M_{kl}(q)\, dq_k\, dq_l \ . \tag{7.7.14}$$

$M_{kl}(q)$ desempenha o papel de um tensor métrico e

$$T = \frac{1}{2}\Big(\frac{d\rho}{dt}\Big)^2 \ , \tag{7.7.15}$$

ou, equivalentemente,

$$dt = \frac{d\rho}{\sqrt{2T}} \ . \tag{7.7.16}$$

Usando este último resultado, o princípio de Maupertuis assume a forma

$$\Delta \int_{\rho_1}^{\rho_2} \sqrt{2T}\, d\rho = 0 \tag{7.7.17}$$

ou, finalmente,

$$\Delta \int_{\rho_1}^{\rho_2} \sqrt{E - V(q)}\, d\rho = 0 \ . \tag{7.7.18}$$

Esta equação costuma ser chamada de *princípio de Jacobi* ou forma de Jacobi do princípio de Maupertuis. No caso de uma só partícula, $d\rho$ é proporcional à distância infinitesimal em três dimensões em coordenadas curvilíneas arbitrárias, e o princípio de Jacobi torna-se idêntico ao princípio de Fermat da óptica geométrica para a propagação da luz num meio não-homogêneo com índice de refração proporcional a $\sqrt{E - V}$. Outros aspectos ainda mais surpreendentes e significativos da analogia entre a mecânica clássica e a óptica geométrica serão explorados no Capítulo 9.

Quando T é constante, o princípio de Maupertuis na forma de Jacobi afirma que o sistema segue a trajetória de menor comprimento no espaço de configuração, isto é, a trajetória física é uma *geodésica* do espaço de configuração. Para uma partícula livre em três dimensões o espaço de configuração confunde-se com o espaço físico tridimensional: as geodésicas são linhas retas, correspondendo ao fato de a partícula descrever um movimento retilíneo uniforme.

■ **Exercício 7.6.1.** Uma partícula move-se livremente no espaço tridimensional, exceto pela restrição de permanecer sobre uma superfície fixa. Mostre que a partícula segue uma geodésica da superfície. ■

PROBLEMAS

7.1. Uma massa m oscila num plano vertical fixo suspensa por um fio de massa desprezível que passa por um orifício numa mesa horizontal. O fio é puxado através do orifício a uma taxa constante, de tal modo que o comprimento pendente no instante t é $r = \ell - \alpha t$, onde ℓ e α são constantes positivas. Escolhendo o ângulo que o fio faz com a vertical como coordenada generalizada, encontre a hamiltoniana do sistema. A hamiltoniana é a energia total? A hamiltoniana é constante de movimento?

7.2. Dada a hamiltoniana $H = p^2/2m + \mathbf{a}\cdot\mathbf{p} + V$, com $\mathbf{a} = \mathbf{a}(\mathbf{r})$ e $V = V(\mathbf{r})$, encontre a lagrangiana correspondente. No caso $V = -\mathbf{F}\cdot\mathbf{r}$ com $\mathbf{a}$ e $\mathbf{F}$ vetores constantes, obtenha a solução das equações de Hamilton com as condições iniciais $\mathbf{r}(0) = 0$ e $\mathbf{p}(0) = 0$.

7.3. Dada a hamiltoniana $H(q_1, q_2, p_1, p_2) = q_1 p_1 - q_2 p_2 - a\, q_1^2 + b\, q_2^2$, prove que

$$F_1 = \frac{p_2 - b q_2}{q_1} \quad , \quad F_2 = q_1 q_2 \quad , \quad F_3 = q_1 e^{-t}$$

são constantes de movimento. Obtenha uma quarta constante de movimento independente destas três e, usando-as, determine a solução das equações de movimento, isto é, $q_1(t), q_2(t), p_1(t), p_2(t)$ em termos de quatro constantes arbitrárias.

7.4. Se a hamiltoniana H e uma quantidade F são constantes de movimento, prove que $\partial F/\partial t$ é constante de movimento. Estenda este resultado para $\partial^n F/\partial t^n$.

7.5. (i) Dada a lagrangiana

$$L = \frac{\dot{Q}^2}{2}\cos^2\omega t - \frac{\omega}{2} Q\dot{Q}\,\mathrm{sen}\,2\omega t - \frac{\omega^2 Q^2}{2}\cos 2\omega t \ ,$$

encontre a hamiltoniana correspondente. (ii) Esta hamiltoniana é constante de movimento? (iii) Mostre que a hamiltoniana expressa em termos da nova variável $q = Q\cos\omega t$ e de seu momento canônico conjugado não depende explicitamente do tempo. Que sistema físico ela representa?

7.6. Considere o problema de N corpos em interação gravitacional do ponto de vista do sistema de referência do centro de massa. A hamiltoniana é dada por $H = T + V$, onde $T = \sum_i |\mathbf{p}_i|^2/2m_i$ e

$$V(\mathbf{r}_1, \ldots, \mathbf{r}_N) = -\frac{1}{2}\sum_{\substack{i,j \\ i\neq j}} \frac{G m_i m_j}{|\mathbf{r}_i - \mathbf{r}_j|} \ .$$

(i) Introduza a quantidade $I = \frac{1}{2}\sum_i m_i |\mathbf{r}_i|^2$ e prove que $\ddot{I} = E + T$. Sugestão: comprove que V é uma função homogênea e aplique o teorema de Euler do Apêndice B. (ii) Levando em conta que T é uma quantidade sempre positiva, prove, fazendo duas integrações sucessivas, que $I(t) \geq I(0) + \dot{I}(0)t + Et^2/2$. Conclua, finalmente, que se a energia total for positiva pelo menos um dos corpos escapará para o infinito no limite $t \to \infty$.

7.7. Considere um planeta descrevendo uma órbita elíptica em torno do Sol, com as coordenadas escolhidas de modo que o plano do movimento tenha $\theta = \pi/2$. Pela Eq.(7.1.15), em coordenadas polares (r, φ) no plano orbital a hamiltoniana escreve-se

$$H = T + V = \frac{1}{2m}\left(p_r^2 + \frac{p_\varphi^2}{r^2}\right) - \frac{k}{r} \ ,$$

com $k > 0$. Aparentemente, o teorema do virial é válido quaisquer que sejam as variáveis canônicas empregadas. Mostre que, em termos das coordenadas polares (r, φ) e de seus momentos conjugados, a Eq.(7.3.3) toma a forma

$$2\langle T\rangle = -\langle V\rangle - \langle \frac{p_\varphi^2}{mr^2}\rangle \ .$$

Comparando este resultado com (7.3.11), mostre que $\langle p_\varphi^2/mr^2\rangle = 0$. Usando o fato de p_φ ser constante de movimento e diferente de zero para órbitas elípticas, conclua que $\langle 1/r^2\rangle = 0$. Argumente que este resultado é impossível. Examinando cuidadosamente a demonstração e as condições de validade do teorema do virial, explique as razões pelas quais o seu emprego em coordenadas polares conduz a um resultado absurdo (Chagas & Lemos 1981).

7.8. O movimento de certas turbulências no hélio líquido é descrito aproximadamente pela hamiltoniana $H(x,p) = A\sqrt{p} - Fx$, A e F constantes. (i) Encontre a solução geral das equações de Hamilton. (ii) Determine a lagrangiana associada a esta hamiltoniana.

7.9. Construa a hamiltoniana associada à lagrangiana do Problema 1.2. H é a energia total? H é constante de movimento?

7.10. Uma partícula de carga e move-se num potencial central $V(r)$ superposto a um campo magnético uniforme $\mathbf{B}$, cujo potencial vetor é $\mathbf{A} = \mathbf{B} \times \mathbf{r}/2$. (i) Mostre que, se $\mathbf{B}$ for um campo fraco, de modo que efeitos proporcionais a B^2 possam ser desprezados,

$$H(\mathbf{r},\mathbf{p}) = \frac{p^2}{2m} + V(\mathbf{r}) - \boldsymbol{\mu}\cdot\mathbf{B} \ ,$$

onde

$$\boldsymbol{\mu} = \frac{e}{2mc}\mathbf{r}\times\mathbf{p} \ .$$

(ii) Escreva as equações de Hamilton nesta aproximação.

7.11. Um sistema físico é descrito pela lagrangiana

$$L = \frac{m}{2}\left(\dot\rho^2 + \rho^2\dot\varphi^2 + \dot z^2\right) + a\rho^2\dot\varphi \ ,$$

onde (ρ, φ, z) são coordenadas cilíndricas e a é uma constante. (i) Determine H. (ii) Identifique três constantes de movimento. (iii) Mostre que a solução da equação radial pode ser reduzida a uma quadratura da forma

$$t = \int \frac{d\rho}{\sqrt{\alpha - (\beta - a\rho^2)^2/m^2\rho^2}} \ ,$$

sendo α e β constantes.

7.12. Partindo da formulação hamiltoniana, obtenha outra formulação da mecânica usando como variáveis $p, \dot p$ e uma função $Y(p, \dot p, t)$. Seguindo o exemplo das Eqs.(7.1.2)

e (7.1.5), a transformação de Legendre que substitui q por $\dot{p}$ deve pressupor que as equações $\dot{p}_i = -\partial H/\partial q_i$ sejam resolvidas para os qs em termos de $p, \dot{p}, t$. A nova função $Y(p, \dot{p}, t)$ deve diferir de $H(q, p, t)$ pela soma dos produtos $q_k \dot{p}_k$ de tal modo que dY só dependa das diferenciais dos ps e $\dot{p}$s. Construa a transformação de Legendre que leva $H(q, p, t)$ a $Y(p, \dot{p}, t)$ e deduza as equações de movimento em termos de Y. Aplique ao caso particular $H = p^2/2 + \omega^2 q^2/2$ e comente os resultados obtidos.

7.13. Encontre a hamiltoniana e as equações de Hamilton para o sistema mecânico do Problema 1.15.

7.14. Na eletrodinâmica de Weber a força entre duas cargas em movimento é gerada pelo potencial generalizado

$$U(r, \dot{r}) = \frac{ee'}{r}\left(1 + \frac{\dot{r}^2}{2c^2}\right) ,$$

onde r é a distância entre as cargas. Considere uma partícula movendo-se num plano sob a ação deste potencial. Usando coordenadas polares, encontre a hamiltoniana $H(r, \theta, p_r, p_\theta)$ e as equações de Hamilton. Mostre que p_θ é constante e reduza o problema radial a uma equação diferencial de primeira ordem para r.

7.15. O movimento de uma partícula num potencial central $V(r)$ é descrito num referencial que gira com velocidade angular constante $\boldsymbol{\omega}$, em relação a um referencial inercial, em torno de um eixo passando pelo centro de força. (i) Mostre que o momento conjugado ao vetor posição $\mathbf{R}$ ($= \mathbf{r}$, mas visto do referencial girante) é $\mathbf{P} = m[\mathbf{V} + \boldsymbol{\omega} \times \mathbf{R}]$, onde $\mathbf{V}$ é a velocidade relativa ao referencial girante. (ii) Construa a função hamiltoniana $H(\mathbf{R}, \mathbf{P}) = \mathbf{P} \cdot \mathbf{V} - L$, onde $L = mv^2/2 - V(r)$. (iii) Mostre que $H(\mathbf{R}, \mathbf{P})$ se conserva mas não coincide com a energia total no referencial girante, isto é, $H \neq mV^2/2 + V(R)$.

7.16. Considere o movimento relativístico de uma partícula num potencial central $V(r)$. (i) Mostre que $\boldsymbol{l} = \mathbf{r} \times \mathbf{p}$ é constante de movimento, onde $\mathbf{p}$ é o momento linear relativístico, e, como conseqüência, a hamiltoniana (7.4.4) pode ser reduzida à forma

$$H = c\sqrt{m^2c^2 + p_r^2 + \frac{p_\varphi^2}{r^2}} + V(r) ,$$

onde r e φ são coordenadas polares no plano da órbita. (ii) Usando as equações de Hamilton, prove que

$$p_r = \frac{p_\varphi}{r^2}\frac{dr}{d\varphi} .$$

(iii) Prove que, em termos da variável $u = 1/r$, a equação da órbita é

$$\varphi = \int \frac{l\, du}{\sqrt{\frac{1}{c^2}[E - V(1/u)]^2 - m^2c^2 - l^2u^2}} ,$$

onde E e l denotam os valores constantes de H e p_φ, respectivamente. (iv) No caso do problema de Kepler relativístico, com $V(r) = -\kappa/r$, mostre que a equação da órbita tem a forma

$$r = \frac{\alpha}{1 + \epsilon\cos\gamma\varphi}$$

com

$$\gamma = \sqrt{1 - \frac{\kappa^2}{c^2 l^2}} \ , \quad \alpha = \frac{c^2 l^2 \gamma^2}{\kappa E} \ , \quad \epsilon = \sqrt{1 + \frac{\gamma^2 (1 - m^2 c^4 / E^2)}{1 - \gamma^2}} \ .$$

(v) No caso de movimento ligado, a órbita não é uma elipse com excentricidade ϵ porque $\gamma \neq 1$, mas comporta-se como uma elipse precessante. No caso em que γ é quase igual a 1, mostre que o ângulo $\Delta\varphi$ entre dois mínimos sucessivos de r é $\Delta\varphi = \pi\kappa^2 / c^2 l^2$.

7.17. Usando o princípio de Maupertuis, mostre que um projétil disparado nas imediações da superfície da Terra segue uma trajetória parabólica.

7.18. O princípio de Maupertuis pode ser estendido a sistemas não-conservativos, isto é, sistemas mecânicos cuja hamiltoniana depende explicitamente do tempo. Com h definida pela equação (2.6.1), mostre que a integral

$$I = \int_{t_1}^{t_2} \left\{ \sum_{i=1}^{n} p_i \dot{q}_i + t \frac{dh}{dt} \right\} dt$$

é estacionária para a trajetória real, se comparada com trajetórias vizinhas com os mesmos pontos terminais e para as quais h tem os mesmos valores terminais.

Capítulo 8

TRANSFORMAÇÕES CANÔNICAS

You boil it in sawdust: you salt it in glue:
You condense it with locusts and tape:
Still keeping one principal object in view –
To preserve its symmetrical shape.
Lewis Carroll, *The Hunting of the Snark*

As equações de Lagrange são invariantes sob uma transformação geral de *coordenadas*, isto é, sua forma permanece a mesma qualquer que seja a escolha das coordenadas generalizadas. Na formulação hamiltoniana as *coordenadas e momentos* são variáveis independentes, o que torna possível considerar mudanças de variáveis *no espaço de fase* que preservem a forma das equações de Hamilton, ampliando enormemente a gama de transformações admissíveis. Essa ampliação, por seu turno, possibilita escolher judiciosamente variáveis canônicas que simplifiquem a hamiltoniana e, em decorrência, facilitem a resolução das equações de movimento.

8.1 Transformações Canônicas e Funções Geradoras

Uma transformação de variáveis no espaço de fase será de nosso interesse apenas se preservar a forma canônica das equações de movimento. Mais precisamente, dadas as variáveis canônicas (q,p), a hamiltoniana $H(q,p,t)$ e as equações de Hamilton

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad , \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \quad , \quad i = 1,\ldots,n \ , \tag{8.1.1}$$

estaremos interessados na transformação inversível

$$Q_i = Q_i(q,p,t) \quad , \quad P_i = P_i(q,p,t) \quad , \quad i = 1,\ldots,n \ , \tag{8.1.2}$$

desde que seja possível encontrar uma função $K(Q,P,t)$ tal que as equações de movimento para as novas variáveis tenham a forma hamiltoniana:

$$\dot{Q}_i = \frac{\partial K}{\partial P_i} \quad , \quad \dot{P}_i = -\frac{\partial K}{\partial Q_i} \quad , \quad i = 1, \ldots, n \ . \tag{8.1.3}$$

Vale destacar que as equações de movimento nas novas variáveis devem ter a forma hamiltoniana *qualquer que seja* a função hamiltoniana original $H(q,p,t)$.

A validade simultânea das Eqs.(8.1.1) e (8.1.3) implica a validade simultânea dos princípios variacionais

$$\delta \int_{t_1}^{t_2} \Big\{ \sum_{i=1}^{n} p_i \, \dot{q}_i \ - \ H(q,p,t) \Big\} \, dt = 0 \ , \tag{8.1.4}$$

e

$$\delta \int_{t_1}^{t_2} \Big\{ \sum_{i=1}^{n} P_i \, \dot{Q}_i \ - \ K(Q,P,t) \Big\} \, dt = 0 \ . \tag{8.1.5}$$

Uma condição *suficiente* para a validade comum de (8.1.4) e (8.1.5) é que os respectivos integrandos difiram pela derivada total em relação ao tempo de uma função $\Phi(q,p,t)$, pois

$$\delta \int_{t_1}^{t_2} \frac{d\Phi}{dt} \, dt = \delta\Phi(q(t_2),p(t_2),t_2) - \delta\Phi(q(t_1),p(t_1),t_1) = 0 \tag{8.1.6}$$

porque $\delta q_i(t_1) = \delta q_i(t_2) = 0$ e $\delta p_i(t_1) = \delta p_i(t_2) = 0$. Assim, a transformação (8.1.2) preserva a forma hamiltoniana das equações de movimento se

$$\sum_{i=1}^{n} p_i \, \dot{q}_i \ - \ H = \sum_{i=1}^{n} P_i \, \dot{Q}_i \ - \ K + \frac{d\Phi}{dt} \ , \tag{8.1.7}$$

ou, em forma diferencial,

$$\sum_{i=1}^{n} \big(p_i \, dq_i \ - P_i \, dQ_i \big) + (K - H)dt = d\Phi \ , \tag{8.1.8}$$

equação que serve para caracterizar uma transformação canônica.

Definição 8.1.1. A transformação inversível (8.1.2) é dita canônica se e somente se existem funções $K(Q,P,t)$ e $\Phi(q,p,t)$ tais que a Eq. (8.1.8) é satisfeita.

A forma de (8.1.8) sugere encarar Φ como função das coordenadas antigas e novas. Suponha que as n primeiras equações (8.1.2) possam ser resolvidas para os n ps em termos de (q,Q,t). Neste caso, as n últimas das Eqs.(8.1.2) permitem escrever os momentos transformados em termos de (q,Q,t). Em outras palavras, podemos tomar (q,Q) como um conjunto de $2n$ variáveis independentes. Definindo a *função geradora* $F_1(q,Q,t)$ por

$$F_1(q,Q,t) = \Phi(q,p(q,Q,t),t) \ , \tag{8.1.9}$$

da Eq.(8.1.8) deduz-se

$$p_i = \frac{\partial F_1}{\partial q_i} \quad , \quad P_i = -\frac{\partial F_1}{\partial Q_i} \quad , \quad i = 1,\ldots,n \ , \tag{8.1.10}$$

e, também,

$$K(Q,P,t) = H(q,p,t) + \frac{\partial F_1}{\partial t} \ . \tag{8.1.11}$$

Note que, dada uma função $F_1(q,Q,t)$, uma transformação canônica fica automaticamente definida pelas Eqs.(8.1.10), com a hamiltoniana transformada fornecida por (8.1.11). De fato, invertendo as n primeiras equações (8.1.10) somos conduzidos a $Q_i = Q_i(q,p,t)$. Introduzindo este resultado no segundo membro de cada uma das n últimas equações (8.1.10) redunda $P_i = P_i(q,p,t)$, e a mudança de variáveis no espaço de fase $(q,p) \longrightarrow (Q,P)$ é canônica por construção.

Pode não ser conveniente ou mesmo não ser possível tomar (q,Q) como variáveis independentes, como no caso da transformação identidade $Q_i = q_i, P_i = p_i$. Mas, se pudermos tomar (q,P) como $2n$ variáveis independentes, existirá uma função geradora do tipo $F_2(q,P,t)$. Com efeito, notando que

$$-\sum_i P_i\, dQ_i \ = \ -d\Big(\sum_i P_i\, Q_i\Big) + \sum_i Q_i\, dP_i \ , \tag{8.1.12}$$

levando esta identidade em (8.1.8) e definindo

$$F_2 = \sum_i P_i\, Q_i + F_1 \ , \tag{8.1.13a}$$

ou, mais precisamente,

$$F_2(q,P,t) = \sum_i P_i\, Q_i(q,P,t) + \Phi(q,p(q,P,t),t) \ , \tag{8.1.13b}$$

a Eq.(8.1.8) assume a forma

$$dF_2 = \sum_{i=1}^{n}\big(p_i\, dq_i \ + Q_i\, dP_i\big) + (K - H)dt \ . \tag{8.1.14}$$

Conseqüentemente,

$$p_i = \frac{\partial F_2}{\partial q_i} \quad , \quad Q_i = \frac{\partial F_2}{\partial P_i} \quad , \quad i = 1,\ldots,n \ , \tag{8.1.15}$$

$$K(Q,P,t) = H(q,p,t) + \frac{\partial F_2}{\partial t} \ . \tag{8.1.16}$$

Como no caso anterior, dada uma função $F_2(q,P,t)$, pela resolução das n primeiras

equações (8.1.15) para P_i obtém-se $P_i = P_i(q,p,t)$, e substituindo este resultado no segundo membro das n últimas Eqs.(8.1.15) resulta $Q_i = Q_i(q,p,t)$, a transformação assim obtida sendo automaticamente canônica, com a hamiltoniana transformada K e a original H relacionadas por (8.1.16).

■ **Exercício 8.1.1.** (i) Suponha que se possa tomar (p,Q) como variáveis independentes. Definindo $F_3(p,Q,t) = -\sum_i q_i p_i + \Phi(q,p,t)$ deduza

$$q_i = -\frac{\partial F_3}{\partial p_i} \quad , \quad P_i = -\frac{\partial F_3}{\partial Q_i} \quad , \quad i = 1,\ldots,n \ , \tag{8.1.17}$$

$$K(Q,P,t) = H(q,p,t) + \frac{\partial F_3}{\partial t} \ . \tag{8.1.18}$$

(ii) Definindo $F_4(p,P,t) = -\sum_i (q_i p_i - Q_i P_i) + \Phi(q,p,t)$, mostre que

$$q_i = -\frac{\partial F_4}{\partial p_i} \quad , \quad Q_i = \frac{\partial F_4}{\partial P_i} \quad , \quad i = 1,\ldots,n \ , \tag{8.1.19}$$

$$K(Q,P,t) = H(q,p,t) + \frac{\partial F_4}{\partial t} \ , \tag{8.1.20}$$

desde que seja possível considerar (p,P) como variáveis independentes. ■

Naturalmente, são possíveis funções geradoras de tipos híbridos, tais como, para um sistema com dois graus de liberdade, $G(q_1,Q_1,q_2,P_2,t)$ que mistura os dois primeiros tipos. Seja como for, funções das variáveis antigas e novas funcionam como "pontes" entre as variáveis antigas e novas estabelecendo uma transformação canônica. Uma vez escolhida uma função geradora, fica definida uma transformação canônica mas não é possível saber *a priori* se tal transformação será de utilidade para simplificar a hamiltoniana de algum sistema físico de interesse: trata-se de uma solução à procura de um problema. Por outro lado, a mera existência de alguma função geradora constitui um critério para determinar se uma mudança de variáveis no espaço de fase é canônica.

■ **Exemplo 8.1.1.** Determinar a transformação canônica gerada por $F_2(q,P,t) = \sum_{k=1}^n q_k P_k$.
Solução. Pelas Eqs.(8.1.15) temos

$$p_i = \frac{\partial F_2}{\partial q_i} = P_i \quad , \quad Q_i = \frac{\partial F_2}{\partial P_i} = q_i \quad , \quad i = 1,\ldots,n \tag{8.1.21}$$

com $K(Q,P,t) = H(q,p,t)$, que é simplesmente a transformação identidade. ■

A existência de uma função geradora caracteriza a transformação identidade como canônica, um resultado esperado. As Eqs.(8.1.21) não permitem que se escolha (q,Q) como um conjunto de $2n$ variáveis independentes, de modo que a transformação identidade não admite uma função geradora do tipo F_1.

■ **Exercício 8.1.2.** Mostre que a transformação canônica gerada por $F_1(q,Q,t) = \sum_{k=1}^n q_k Q_k$ é

$$p_i = Q_i \quad , \quad P_i = -q_i \quad , \quad i = 1,\ldots,n \ , \tag{8.1.22}$$

que, salvo por um sinal, constitui um intercâmbio entre coordenadas e momentos. Esta transformação admite uma função geradora do tipo F_2? ■

■ **Exemplo 8.1.2.** Obter a transformação canônica gerada por $F_1(q,Q,t) = m(q-Q)^2/2t$ e aplicá-la à resolução do problema da partícula livre unidimensional, cuja hamiltoniana é $H = p^2/2m$.

Solução. Pelas Eqs.(8.1.10) temos

$$p = \frac{\partial F_1}{\partial q} = \frac{m(q-Q)}{t} \quad , \quad P = -\frac{\partial F_1}{\partial Q} = \frac{m(q-Q)}{t} \ , \tag{8.1.23}$$

ou, em forma direta,

$$Q = q - \frac{pt}{m} \quad , \quad P = p \ . \tag{8.1.24}$$

Por outro lado,

$$K(Q,P,t) = H(q,p,t) + \frac{\partial F_1}{\partial t} = \frac{p^2}{2m} - \frac{m(q-Q)^2}{2t^2} = \frac{P^2}{2m} - \frac{P^2}{2m} = 0 \ , \tag{8.1.25}$$

e as equações de Hamilton transformadas são trivialmente resolvidas:

$$\dot{Q} = \frac{\partial K}{\partial P} = 0 \implies Q = a \quad ; \quad \dot{P} = -\frac{\partial K}{\partial Q} = 0 \implies P = b \ , \tag{8.1.26}$$

onde a e b são constantes arbitrárias. Retornando às variáveis canônicas originais resulta, finalmente,

$$q = a + \frac{b}{m}t \ , \tag{8.1.27}$$

que é a solução geral da equação de movimento da partícula livre. ■

■ **Exemplo 8.1.3.** Completar a mudança de variáveis no espaço de fase (a e b são constantes)

$$P = \frac{1}{b^2}(p^2 + a^2q^2) \tag{8.1.28}$$

de modo a torná-la uma transformação canônica. Aplicá-la, em seguida, à resolução do problema do oscilador harmônico unidimensional, cuja hamiltoniana é $H = p^2/2m + m\omega^2q^2/2$.

Solução. Precisamos encontrar $Q = Q(q,p)$ de modo que, juntamente com (8.1.28), fique definida uma transformação canônica. Escrevendo

$$p = (b^2P - a^2q^2)^{1/2} = b(P - c^2q^2)^{1/2} \quad , \quad c = \frac{a}{b} \ , \tag{8.1.29}$$

e comparando com (8.1.15), concluímos ser adequado buscar uma função geradora do tipo $F_2(q,P)$. Devemos ter, portanto

$$p = \frac{\partial F_2}{\partial q} = b(P - c^2q^2)^{1/2} \implies F_2 = b\int \sqrt{P - c^2q^2}\, dq \ . \tag{8.1.30}$$

Esta integral é elementar, porém é mais conveniente não efetuá-la, pois estamos na verdade interessados na outra metade da transformação canônica, dada por

$$Q = \frac{\partial F_2}{\partial P} = \frac{b}{2}\int \frac{dq}{(P - c^2q^2)^{1/2}} = \frac{b}{2c}\mathrm{sen}^{-1}\left(\frac{cq}{\sqrt{P}}\right) = \frac{b^2}{2a}\mathrm{sen}^{-1}\left(\frac{aq}{b\sqrt{P}}\right) \ , \tag{8.1.31}$$

onde descartamos uma "constante" de integração $C(P)$ a fim de ficar com a transformação canônica mais simples possível. Assim,

$$Q = \frac{b^2}{2a}\text{sen}^{-1}\left\{\frac{aq}{(p^2+a^2q^2)^{1/2}}\right\} \ . \tag{8.1.32}$$

A transformação inversa tem a forma

$$q = \frac{b}{a}\sqrt{P}\,\text{sen}\left(\frac{2aQ}{b^2}\right) \quad , \quad p = b\sqrt{P}\cos\left(\frac{2aQ}{b^2}\right) \ . \tag{8.1.33}$$

A hamiltoniana do oscilador pode ser escrita na forma

$$H(q,p) = \frac{1}{2m}(p^2 + m^2\omega^2 q^2) \ , \tag{8.1.34}$$

o que sugere escolher $a = m\omega$ em (8.1.28) e, para simplificar as Eqs.(8.1.33), $b = \sqrt{2a} = \sqrt{2m\omega}$. Assim procedendo, como F_2 não depende explicitamente do tempo, resulta

$$K(Q,P) = H(q,p) = \frac{p^2}{2m} + \frac{m\omega^2}{2}q^2 = \frac{\omega}{2m\omega}(p^2 + m^2\omega^2 q^2) = \omega P \ , \tag{8.1.35}$$

de modo que

$$\dot{Q} = \frac{\partial K}{\partial P} = \omega \quad \Longrightarrow \quad Q = \omega t + \delta \ , \tag{8.1.36a}$$

$$\dot{P} = -\frac{\partial K}{\partial Q} = 0 \quad \Longrightarrow \quad P = \text{constante} = \frac{E}{\omega} \ , \tag{8.1.36b}$$

onde E denota a energia total do oscilador (valor constante de H). Retornando à variável original q com a ajuda de (8.1.33), obtém-se, finalmente,

$$q = \sqrt{\frac{2E}{m\omega^2}}\,\text{sen}\,(\omega t + \delta) \ , \tag{8.1.37}$$

que é a solução usual da equação de movimento do oscilador. ■

Somente um insensato recorreria a transformações canônicas para resolver os problemas da partícula livre ou do oscilador harmônico simples, sistemas cujas equações de movimento podem ser trivialmente resolvidas de forma direta. O objetivo dos dois últimos exemplos foi ilustrar, em casos elementares, como uma transformação canônica cuidadosamente escolhida é capaz de simplificar drasticamente a forma de uma dada hamiltoniana, podendo até mesmo reduzi-la a zero. A elaboração de um esquema sistemático de construção de uma transformação canônica capaz de anular a hamiltoniana transformada constitui a essência do método de Hamilton-Jacobi, a ser discutido no próximo capítulo.

■ **Exercício 8.1.3.** Exibindo uma função geradora, mostre que a *transformação de ponto* $Q_i = f_i(q_1, \ldots, q_n, t)$ é canônica e obtenha a expressão dos momentos transformados em termos dos momentos originais. ■

8.2 Canonicidade e Parênteses de Lagrange

A validade de (8.1.8) para t variável implica sua validade para t fixo, isto é

$$\sum_{i=1}^{n}\Big[\, p_i\, dq_i \;-\; P_i(q,p,t)\, dQ_i(q,p,t)\,\Big] = d\Phi(q,p,t) \qquad (t \text{ fixo}) \;, \tag{8.2.1}$$

onde as diferenciais só levam em conta as variações dos qs e ps, com o tempo mantido constante. Reciprocamente, a validade desta equação com t fixo implica a validade de (8.1.8) para t variável (Problema 8.1). O teste do caráter canônico de uma transformação por meio de funções geradoras não é direto, pois exige a inversão de metade das equações (8.1.2) para obter um conjunto de $2n$ variáveis independentes, constituído por n variáveis novas e n variáveis antigas, seguida da exibição explícita de uma função geradora. A Eq.(8.2.1) permite obter um critério direto e geral de canonicidade em termos dos chamados *colchetes* ou *parênteses de Lagrange*, que passamos a definir.

Definição 8.2.1. Seja (η,ξ) um conjunto de $2n$ variáveis canônicas dependentes dos parâmetros u, v. O colchete ou parêntese de Lagrange de u e v em relação a (η,ξ), denotado por $[u,v]_{(\eta,\xi)}$, é definido por

$$[u,v]_{(\eta,\xi)} = \sum_{k=1}^{n}\left(\frac{\partial\eta_k}{\partial u}\,\frac{\partial\xi_k}{\partial v} - \frac{\partial\eta_k}{\partial v}\,\frac{\partial\xi_k}{\partial u}\right) . \tag{8.2.2}$$

O próximo resultado estabelece a utilidade dos parênteses de Lagrange para caracterizar uma transformação como canônica.

Teorema 8.2.1. A transformação (8.1.2) é canônica se e somente se

$$[q_i,q_j]_{(Q,P)} = 0 \quad , \quad [p_i,p_j]_{(Q,P)} = 0 \quad , \quad [q_i,p_j]_{(Q,P)} = \delta_{ij} \;. \tag{8.2.3}$$

Demonstração. Como já vimos, a transformação 8.1.2 é canônica se e somente se

$$\sum_{j=1}^{n}\left[\, p_j\, dq_j \;-\; P_j(q,p,t)\sum_{k=1}^{n}\Big(\frac{\partial Q_j}{\partial q_k}\, dq_k + \frac{\partial Q_j}{\partial p_k}\, dp_k\Big)\right] = d\Phi(q,p,t) \qquad (t \text{ fixo}) \;, \tag{8.2.4}$$

ou seja,

$$\sum_{j=1}^{n}\big(A_j\, dq_j \;+\; B_j\, dp_j\big) = d\Phi(q,p,t) \qquad (t \text{ fixo}) \;, \tag{8.2.5}$$

onde

$$A_j = p_j - \sum_{l=1}^{n} P_l\,\frac{\partial Q_l}{\partial q_j} \quad , \quad B_j = -\sum_{l=1}^{n} P_l\,\frac{\partial Q_l}{\partial p_j} \;. \tag{8.2.6}$$

Como está provado no Apêndice D, a condição necessária e suficiente para que o primeiro

esquerdo da Eq.(8.2.5) seja uma diferencial exata é a igualdade de todas as derivadas cruzadas dos coeficientes das diferenciais dq_j e dp_j:

$$\frac{\partial A_j}{\partial q_i} = \frac{\partial A_i}{\partial q_j} \quad , \quad \frac{\partial B_j}{\partial p_i} = \frac{\partial B_i}{\partial p_j} \quad , \quad \frac{\partial A_i}{\partial p_j} = \frac{\partial B_j}{\partial q_i} \ . \tag{8.2.7}$$

A primeira destas equações fornece

$$-\sum_{l=1}^{n}\Big(\frac{\partial P_l}{\partial q_i}\frac{\partial Q_l}{\partial q_j} + P_l\frac{\partial^2 Q_l}{\partial q_i \partial q_j}\Big) = -\sum_{l=1}^{n}\Big(\frac{\partial P_l}{\partial q_j}\frac{\partial Q_l}{\partial q_i} + P_l\frac{\partial^2 Q_l}{\partial q_j \partial q_i}\Big) \ , \tag{8.2.8}$$

donde

$$[q_i, q_j]_{(Q,P)} = \sum_{l=1}^{n}\Big(\frac{\partial Q_l}{\partial q_i}\frac{\partial P_l}{\partial q_j} - \frac{\partial Q_l}{\partial q_j}\frac{\partial P_l}{\partial q_i}\Big) = 0 \ . \tag{8.2.9}$$

De modo análogo, a segunda das Eqs.(8.2.7) implica

$$[p_i, p_j]_{(Q,P)} = \sum_{l=1}^{n}\Big(\frac{\partial Q_l}{\partial p_i}\frac{\partial P_l}{\partial p_j} - \frac{\partial Q_l}{\partial p_j}\frac{\partial P_l}{\partial p_i}\Big) = 0 \ . \tag{8.2.10}$$

Finalmente, a terceira das Eqs.(8.2.7) equivale a

$$\delta_{ij} - \sum_{l=1}^{n}\Big(\frac{\partial P_l}{\partial p_j}\frac{\partial Q_l}{\partial q_i} + P_l\frac{\partial^2 Q_l}{\partial p_j \partial q_i}\Big) = -\sum_{l=1}^{n}\Big(\frac{\partial P_l}{\partial q_i}\frac{\partial Q_l}{\partial p_j} + P_l\frac{\partial^2 Q_l}{\partial q_i \partial p_j}\Big) \ , \tag{8.2.11}$$

ou seja,

$$[q_i, p_j]_{(Q,P)} = \sum_{l=1}^{n}\Big(\frac{\partial Q_l}{\partial q_i}\frac{\partial P_l}{\partial p_j} - \frac{\partial Q_l}{\partial p_j}\frac{\partial P_l}{\partial q_i}\Big) = \delta_{ij} \ , \tag{8.2.12}$$

completando a demonstração do teorema. □

■ **Exemplo 8.2.1.** Provar que a transformação

$$Q_1 = p_1^2 \quad , \quad Q_2 = p_2^2 + q_2 \quad , \quad P_1 = -\frac{q_1}{2p_1} \quad , \quad P_2 = p_2 \tag{8.2.13}$$

é canônica.

Solução. Temos que

$$[q_1, q_2]_{(Q,P)} = \Big(\frac{\partial Q_1}{\partial q_1}\frac{\partial P_1}{\partial q_2} - \frac{\partial Q_1}{\partial q_2}\frac{\partial P_1}{\partial q_1}\Big) + \Big(\frac{\partial Q_2}{\partial q_1}\frac{\partial P_2}{\partial q_2} - \frac{\partial Q_2}{\partial q_2}\frac{\partial P_2}{\partial q_1}\Big) = 0 \ ,$$

$$[p_1, p_2]_{(Q,P)} = \Big(\frac{\partial Q_1}{\partial p_1}\frac{\partial P_1}{\partial p_2} - \frac{\partial Q_1}{\partial p_2}\frac{\partial P_1}{\partial p_1}\Big) + \Big(\frac{\partial Q_2}{\partial p_1}\frac{\partial P_2}{\partial p_2} - \frac{\partial Q_2}{\partial p_2}\frac{\partial P_2}{\partial p_1}\Big) = 0 \ ,$$

com cálculos análogos mostrando que $[q_1, p_2]_{(Q,P)} = [q_2, p_1]_{(Q,P)} = 0$. Finalmente,

$$[q_1, p_1]_{(Q,P)} = \Big(\frac{\partial Q_1}{\partial q_1}\frac{\partial P_1}{\partial p_1} - \frac{\partial Q_1}{\partial p_1}\frac{\partial P_1}{\partial q_1}\Big) + \Big(\frac{\partial Q_2}{\partial q_1}\frac{\partial P_2}{\partial p_1} - \frac{\partial Q_2}{\partial p_1}\frac{\partial P_2}{\partial q_1}\Big) = -2p_1\Big(-\frac{1}{2p_1}\Big) = 1$$

e, ainda,

$$[q_2, p_2]_{(Q,P)} = \left(\frac{\partial Q_1}{\partial q_2}\frac{\partial P_1}{\partial p_2} - \frac{\partial Q_1}{\partial p_2}\frac{\partial P_1}{\partial q_2}\right) + \left(\frac{\partial Q_2}{\partial q_2}\frac{\partial P_2}{\partial p_2} - \frac{\partial Q_2}{\partial p_2}\frac{\partial P_2}{\partial q_2}\right) = 1 \ .$$

É desnecessário calcular os demais parênteses de Lagrange por causa de sua anti-simetria, isto é, $[u, v]_{(\eta,\xi)} = -[v, u]_{(\eta,\xi)}$. Uma vez que todas as condições (8.2.3) são satisfeitas, a transformação é canônica. ■

8.3 Notação Simplética

É possível introduzir uma notação compacta que resume todas as equações de Hamilton numa única equação matricial. Essa notação é particularmente útil para simplificar certas fórmulas e demonstrações que na notação tradicional são bastante complicadas. Seja $\mathbf{z}$ uma matriz-coluna com $2n$ elementos $z_1, \ldots, z_{2n}$, definidos por

$$z_i = q_i \quad , \quad z_{n+i} = p_i \quad , \quad i = 1, \ldots, n \ . \tag{8.3.1}$$

Em palavras: os n primeiros zs são os n qs e os n últimos zs são os n ps. Analogamente, a matriz $\partial H/\partial \mathbf{z}$ é uma coluna com os seguintes elementos:

$$\left(\frac{\partial H}{\partial \mathbf{z}}\right)_i = \frac{\partial H}{\partial q_i} \quad , \quad \left(\frac{\partial H}{\partial \mathbf{z}}\right)_{n+i} = \frac{\partial H}{\partial p_i} \quad , \quad i = 1, \ldots, n \ . \tag{8.3.2}$$

Seja, finalmente, $\mathbf{J}$ a matriz $(2n) \times (2n)$ definida por

$$\mathbf{J} = \begin{pmatrix} \mathbf{O} & I \\ -I & \mathbf{O} \end{pmatrix} \tag{8.3.3}$$

onde $\mathbf{O}$ é a matriz $n \times n$ com todos os elementos nulos e I é a matriz identidade $n \times n$. As equações de Hamilton (7.1.9) podem ser expressas sucintamente como

$$\dot{\mathbf{z}} = \mathbf{J}\frac{\partial H}{\partial \mathbf{z}} \ , \tag{8.3.4}$$

conhecida como *forma simplética*[1] das equações de Hamilton.

■ **Exercício 8.3.1.** Mostre que (8.3.4) equivale às equações de Hamilton na forma usual. ■

A matriz $\mathbf{J}$ possui as seguintes propriedades, cuja verificação é deixada a cargo do leitor:

(1) $\mathbf{J}^2 = -I$, onde I é a matriz identidade $(2n) \times (2n)$;
(2) $\mathbf{J}^T\mathbf{J} = \mathbf{J}\mathbf{J}^T = I \implies \mathbf{J}^{-1} = \mathbf{J}^T = -\mathbf{J}$ ($\mathbf{J}$ é uma matriz ortogonal);
(3) $\det \mathbf{J} = 1$.

[1] Esta palavra foi introduzida pelo matemático Hermann Weyl em 1939 e deriva de uma raiz grega que significa "entremeado" ou "entrelaçado".

Com um agrupamento análogo das variáveis transformadas (Q, P) para formar a matriz-coluna $\boldsymbol{\zeta}$ com $2n$ linhas, uma transformação no espaço de fase escreve-se

$$\boldsymbol{\zeta} = \boldsymbol{\zeta}(\mathbf{z}, t) \tag{8.3.5}$$

ou, em componentes,

$$\zeta_k = \zeta_k(z_1, \ldots, z_{2n}, t) \quad , \quad k = 1, \ldots, 2n \ . \tag{8.3.6}$$

Com esta notação podemos escrever

$$[u, v]_{\mathbf{z}} = \left(\frac{\partial \mathbf{z}}{\partial u}\right)^T \mathbf{J} \frac{\partial \mathbf{z}}{\partial v} \equiv \sum_{r,s=1}^{2n} \frac{\partial z_r}{\partial u} J_{rs} \frac{\partial z_s}{\partial v} \ . \tag{8.3.7}$$

■ **Exercício 8.2.2.** Comprove que (8.3.7) é equivalente ao parêntese de Lagrange $[u, v]_{(q,p)}$ conforme a Definição 8.2.1. ■

Lema 8.3.1. Em notação simplética, as condições de canonicidade (8.2.3) podem ser postas na forma concisa $[z_r, z_s]_{\boldsymbol{\zeta}} = J_{rs}$.

Demonstração. Se $r, s \in \{1, \ldots, n\}$ ou $r, s \in \{n+1, \ldots, 2n\}$ o primeiro membro de $[z_r, z_s]_{\boldsymbol{\zeta}} = J_{rs}$ reduz-se aos parênteses de Lagrange dos qs entre si ou dos ps entre si e o segundo membro é nulo, ao passo que se $r = i \in \{1, \ldots, n\}$ e $s = (n+j) \in \{n+1, \ldots, 2n\}$ a equação $[z_r, z_s]_{\boldsymbol{\zeta}} = J_{rs}$ reduz-se a $[q_i, p_j]_{(Q,P)} = J_{i,n+j} = \delta_{ij}$. A inversão dos papéis de r, s repete este último resultado devido à anti-simetria de J_{rs} . □

Seja, agora, $\mathbf{M}$ a matriz jacobiana da transformação (8.3.5), isto é,

$$M_{rs} = \frac{\partial \zeta_r}{\partial z_s} \ . \tag{8.3.8}$$

Podemos escrever

$$[z_r, z_s]_{\boldsymbol{\zeta}} = \sum_{k,l=1}^{2n} \frac{\partial \zeta_k}{\partial z_r} J_{kl} \frac{\partial \zeta_l}{\partial z_s} = \sum_{k,l=1}^{2n} M_{kr} J_{kl} M_{ls} \ , \tag{8.3.9}$$

ou, apelando para a definição de produto de matrizes,

$$[z_r, z_s]_{\boldsymbol{\zeta}} = \left(\mathbf{M}^T \mathbf{J} \mathbf{M}\right)_{rs} \ . \tag{8.3.10}$$

Teorema 8.3.1. A transformação (8.3.5) é canônica se e somente se sua matriz jacobiana satisfaz

$$\mathbf{M}^T \mathbf{J} \mathbf{M} = \mathbf{J} \ . \tag{8.3.11}$$

Demonstração. Combine (8.3.10) com o Lema 8.3.1 e o Teorema 8.2.1. □

Qualquer matriz que satisfaz a *condição simplética* (8.3.11) é dita uma *matriz simplética*. Cumpre chamar a atenção para a semelhança entre a Eq.(8.3.11) e a condição (6.1.20) que caracteriza uma transformação de Lorentz. A condição simplética pode ser utilizada para testar diretamente se uma dada transformação é canônica.

■ **Exemplo 8.2.2.** Provar que a transformação $Q = (q-p)/\sqrt{2}$, $P = (q+p)/\sqrt{2}$ é canônica.
Solução. Temos que

$$\mathbf{M}^T\mathbf{J}\mathbf{M} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \mathbf{J}\ ,$$

comprovando a natureza canônica da transformação. ■

Teorema 8.3.2. O determinante da matriz jacobiana de uma transformação canônica é igual a ± 1.

■ **Exercício 8.3.2.** Tomando o determinante de (8.3.11), prove o Teorema 8.3.2. ■

É possível ir mais além e demonstrar que o determinante da matriz jacobiana de qualquer transformação canônica é igual a $+1$ (Landau & Lifchitz 1966), mas não necessitaremos desse resultado mais refinado.

Invariância dos Parênteses de Lagrange

A notação simplética simplifica substancialmente a demonstração de uma importante propriedade de invariância dos parênteses de Lagrange.

Teorema 8.3.3. Os parênteses de Lagrange são invariantes sob transformações canônicas, isto é, se a transformação $\mathbf{z} \to \boldsymbol{\zeta}$ é canônica então

$$[u,v]_{\mathbf{z}} = [u,v]_{\boldsymbol{\zeta}}\ . \tag{8.3.12}$$

Demonstração. Note inicialmente que, pela regra da cadeia,

$$\frac{\partial \zeta_r}{\partial u} = \sum_{s=1}^{2n} \frac{\partial \zeta_r}{\partial z_s}\frac{\partial z_s}{\partial u} = \sum_{s=1}^{2n} M_{rs}\frac{\partial z_s}{\partial u} \tag{8.3.13}$$

ou, em notação matricial,

$$\frac{\partial \boldsymbol{\zeta}}{\partial u} = \mathbf{M}\frac{\partial \mathbf{z}}{\partial u}\ . \tag{8.3.14}$$

Levando este resultado em (8.3.7) somos conduzidos a

$$[u,v]_{\boldsymbol{\zeta}} = \left(\frac{\partial \boldsymbol{\zeta}}{\partial u}\right)^T \mathbf{J}\frac{\partial \boldsymbol{\zeta}}{\partial v} = \left(\frac{\partial \mathbf{z}}{\partial u}\right)^T \mathbf{M}^T\mathbf{J}\mathbf{M}\frac{\partial \mathbf{z}}{\partial v} = \left(\frac{\partial \mathbf{z}}{\partial u}\right)^T \mathbf{J}\frac{\partial \mathbf{z}}{\partial v} = [u,v]_{\mathbf{z}}\ , \tag{8.3.15}$$

onde usamos (8.3.11). □

■ **Exercício 8.3.3.** Usando $\mathbb{J}^2 = -I$ e o fato de a inversa de uma matriz pela esquerda coincidir com a inversa pela direita, mostre que

$$\mathbf{M}\mathbb{J}\mathbf{M}^T = \mathbb{J} \tag{8.3.16}$$

é uma conseqüência de (8.3.11). ■

Das Eqs.(8.3.11) e (8.3.16) deduz-se facilmente que as transformações canônicas formam um grupo: (i) a transformação identidade é canônica; (ii) o produto de transformações canônicas é associativo; (iii) duas transformações canônicas sucessivas constituem uma transformação canônica; (iv) a inversa de uma transformação canônica é uma transformação canônica.

■ **Exercício 8.3.4.** Prove que as transformações canônicas formam um grupo. ■

O conjunto de todas as matrizes $(2n)\times(2n)$ que satisfazem (8.3.11) constitui o *grupo real simplético* $Sp_{2n}(\boldsymbol{R})$ sobre o espaço $\boldsymbol{R}^{2n}$.

8.4 Parênteses de Poisson

Seja $F(q,p,t)$ uma *variável dinâmica* arbitrária, isto é, uma função qualquer das variáveis canônicas e do tempo. Então

$$\frac{dF}{dt} = \sum_{k=1}^{n}\Big(\frac{\partial F}{\partial q_k}\dot{q}_k + \frac{\partial F}{\partial p_k}\dot{p}_k\Big) + \frac{\partial F}{\partial t} = \sum_{k=1}^{n}\Big(\frac{\partial F}{\partial q_k}\frac{\partial H}{\partial p_k} - \frac{\partial F}{\partial p_k}\frac{\partial H}{\partial q_k}\Big) + \frac{\partial F}{\partial t}\ , \tag{8.4.1}$$

tendo sido usadas as equações de Hamilton. O *colchete* ou *parêntese de Poisson* $\{F,G\}$ de duas variáveis dinâmicas F e G é definido por

$$\{F,G\} = \sum_{k=1}^{n}\Big(\frac{\partial F}{\partial q_k}\frac{\partial G}{\partial p_k} - \frac{\partial F}{\partial p_k}\frac{\partial G}{\partial q_k}\Big)\ , \tag{8.4.2}$$

de modo que

$$\frac{dF}{dt} = \{F,H\} + \frac{\partial F}{\partial t}\ . \tag{8.4.3}$$

Em particular, tomando primeiro $F = q_i$ e em seguida $F = p_i$, resultam as equações de Hamilton em termos de parênteses de Poisson:

$$\dot{q}_i = \{q_i,H\} \quad , \quad \dot{p}_i = \{p_i,H\}\ , \tag{8.4.4}$$

porque, para ambas as escolhas, F não apresenta dependência temporal *explícita*.

Invariância dos Parênteses de Poisson

Uma das principais vantagens para análises teóricas de se escrever a equação de movimento de uma variável dinâmica arbitrária na forma (8.4.3) reside na propriedade do parêntese de Poisson de ser invariante sob transformações canônicas. Em outras palavras, a Eq.(8.4.3) independe do conjunto de variáveis canônicas escolhido para descrever a dinâmica. Para abreviar a demonstração da invariância do parêntese de Poisson é conveniente reescrevê-lo em notação simplética na forma

$$\{F,G\}_{\mathbf{z}} = \left(\frac{\partial F}{\partial \mathbf{z}}\right)^T \mathbf{J} \frac{\partial G}{\partial \mathbf{z}} \equiv \sum_{r,s=1}^{2n} \frac{\partial F}{\partial z_r} J_{rs} \frac{\partial G}{\partial z_s} \ , \tag{8.4.5}$$

o subscrito indicando cálculo relativamente às variáveis canônicas $\mathbf{z} = (q,p)$.

Teorema 8.4.1. O parêntese de Poisson é invariante sob transformações canônicas, isto é,

$$\{F,G\}_{\mathbf{z}} = \{F,G\}_{\boldsymbol{\zeta}} \tag{8.4.6}$$

se a transformação $\mathbf{z} \to \boldsymbol{\zeta}$ é canônica.

Demonstração. De acordo com a regra da cadeia,

$$\frac{\partial F}{\partial z_r} = \sum_{s=1}^{2n} \frac{\partial \zeta_s}{\partial z_r} \frac{\partial F}{\partial \zeta_s} \qquad \Longrightarrow \qquad \frac{\partial F}{\partial \mathbf{z}} = \mathbf{M}^T \frac{\partial F}{\partial \boldsymbol{\zeta}} \ . \tag{8.4.7}$$

Conseqüentemente,

$$\{F,G\}_{\mathbf{z}} = \left(\frac{\partial F}{\partial \mathbf{z}}\right)^T \mathbf{J} \frac{\partial G}{\partial \mathbf{z}} = \left(\frac{\partial F}{\partial \boldsymbol{\zeta}}\right)^T \mathbf{M}\mathbf{J}\mathbf{M}^T \frac{\partial G}{\partial \boldsymbol{\zeta}} = \left(\frac{\partial F}{\partial \boldsymbol{\zeta}}\right)^T \mathbf{J} \frac{\partial G}{\partial \boldsymbol{\zeta}} = \{F,G\}_{\boldsymbol{\zeta}} \ , \tag{8.4.8}$$

onde usamos a Eq.(8.3.16). □

Canonicidade e Parênteses de Poisson

Como o leitor talvez já tenha suspeitado, os parênteses de Lagrange e de Poisson satisfazem uma importante relação de reciprocidade.

Teorema 8.4.2. Seja $u_r = u_r(z_1,\ldots,z_{2n})$, $r = 1,\ldots,2n$, um conjunto de $2n$ funções mutuamente independentes, de modo que, reciprocamente, $z_r = z_r(u_1,\ldots,u_{2n})$. Então

$$\sum_{k=1}^{2n} \{u_r,u_k\}_{\mathbf{z}} [u_k,u_s]_{\mathbf{z}} = -\delta_{rs} \ . \tag{8.4.9}$$

Demonstração. Segundo (8.3.7) e (8.4.5),

$$\Gamma_{rs} \equiv \sum_{k=1}^{2n} \{u_r,u_k\}_{\mathbf{z}} [u_k,u_s]_{\mathbf{z}} = \sum_{k=1}^{2n} \sum_{l,m=1}^{2n} \sum_{i,j=1}^{2n} \frac{\partial u_r}{\partial z_l} J_{lm} \frac{\partial u_k}{\partial z_m} \frac{\partial z_i}{\partial u_k} J_{ij} \frac{\partial z_j}{\partial u_s}$$

$$= \sum_{l,m=1}^{2n} \sum_{i,j=1}^{2n} \frac{\partial u_r}{\partial z_l} J_{lm} \frac{\partial z_i}{\partial z_m} J_{ij} \frac{\partial z_j}{\partial u_s} = \sum_{l,m=1}^{2n} \sum_{j=1}^{2n} \frac{\partial u_r}{\partial z_l} J_{lm} J_{mj} \frac{\partial z_j}{\partial u_s} \ ,$$

onde usamos a regra da cadeia da diferenciação e $\partial z_i / \partial z_m = \delta_{im}$. Por outro lado,

$$\sum_{m=1}^{2n} J_{lm} J_{mj} = \left(\mathbf{J}^2\right)_{lj} = \left(-\boldsymbol{I}\right)_{lj} = -\delta_{lj} \ ,$$

donde

$$\Gamma_{rs} = -\sum_{l,j=1}^{2n} \frac{\partial u_r}{\partial z_l} \delta_{lj} \frac{\partial z_j}{\partial u_s} = -\sum_{l=1}^{2n} \frac{\partial u_r}{\partial z_l} \frac{\partial z_l}{\partial u_s} = -\frac{\partial u_r}{\partial u_s} = -\delta_{rs}$$

pois os us são mutuamente independentes por hipótese, e a demonstração do teorema está completa. □

Em virtude deste resultado, é possível reformular o critério básico de canonicidade para transformações no espaço de fase em termos de parênteses de Poisson.

Teorema 8.4.3. A transformação $\mathbf{z} \to \boldsymbol{\zeta}$ é canônica se e somente se

$$\{\zeta_r, \zeta_s\}_{\mathbf{z}} = J_{rs} \ , \tag{8.4.10a}$$

ou, em notação tradicional,

$$\{Q_i, Q_j\}_{(q,p)} = 0 \quad , \quad \{P_i, P_j\}_{(q,p)} = 0 \quad , \quad \{Q_i, P_j\}_{(q,p)} = \delta_{ij} \ . \tag{8.4.10b}$$

Demonstração. Sejam $\mathbf{X}$ e $\mathbf{Y}$ matrizes cujos elementos são $\{\zeta_r, \zeta_s\}_{\mathbf{z}}$ e $[\zeta_r, \zeta_s]_{\mathbf{z}}$, respectivamente. Em linguagem matricial a Eq.(8.4.9) equivale a $\mathbf{XY} = -\boldsymbol{I}$. Como a inversa de uma transformação canônica é também uma transformação canônica, o Teorema 8.2.1, com os papéis das variáveis novas e antigas intercambiados, afirma que a transformação (8.3.5) é canônica se e somente se $\mathbf{Y} = \mathbf{J}$. Mas, como $\mathbf{J}^2 = -\boldsymbol{I}$, segue-se imediatamente que $\mathbf{X} = \mathbf{J}$ se e somente se $\mathbf{Y} = \mathbf{J}$, completando a demonstração do teorema. □

■ **Exercício 8.4.1.** Considere um sistema mecânico com um único grau de liberdade. Usando o critério dos parênteses de Poisson, prove que a transformação

$$Q = \ln\left(\frac{\operatorname{sen} p}{q}\right) \quad , \quad P = q \cot p \tag{8.4.11}$$

é canônica. ■

Propriedades Algébricas dos Parênteses de Poisson

Os parênteses de Poisson gozam das seguintes propriedades algébricas:

(P1) Anti-simetria: $\{A,B\} = -\{B,A\}$ donde $\{A,A\} = 0$.
(P2) Linearidade: $\{A+\alpha B, C\} = \{A,C\} + \alpha\{B,C\}$, α independente de (q,p).
(P3) $\{AB,C\} = A\{B,C\} + \{A,C\}B$; $\{A,BC\} = \{A,B\}C + B\{A,C\}$.
(P4) Identidade de Jacobi: $\{\{A,B\},C\} + \{\{B,C\},A\} + \{\{C,A\},B\} = 0$.

Exceto a identidade de Jacobi, cuja demonstração direta requer um longo trabalho algébrico, as demais propriedades são de demonstração fácil e rápida, ficando como exercício para o leitor. Reservamos para a próxima seção uma prova curta e virtualmente sem cálculos da identidade de Jacobi. Uma propriedade adicional que merece menção é

$$\frac{\partial}{\partial\lambda}\{A,B\} = \{\frac{\partial A}{\partial\lambda}, B\} + \{A, \frac{\partial B}{\partial\lambda}\} \ , \tag{8.4.12}$$

onde λ é um dos qs, um dos ps, o tempo t ou um outro parâmetro qualquer.

Os *parênteses de Poisson fundamentais*

$$\{q_i, q_j\} = 0 \quad , \quad \{p_i, p_j\} = 0 \quad , \quad \{q_i, p_j\} = \delta_{ij} \ , \tag{8.4.13}$$

juntamente com as propriedades P1 a P4, permitem, em muitos casos, o cálculo de parênteses de Poisson por métodos puramente algébricos, sem a necessidade de recurso à definição (8.4.2). Por outro lado, como a análise anterior demonstrou, a invariância dos parênteses de Poisson fundamentais é condição necessária e suficiente para que uma transformação no espaço de fase seja canônica.

Uma álgebra de Lie consiste num espaço vetorial de elementos $X, Y, \ldots$ no qual está definido um "produto" bilinear $[X,Y]$ que satisfaz a propriedade de anti-simetria e obedece à identidade de Jacobi (Arnold 1976; Sternberg 1994). Assim, as funções infinitamente diferenciáveis sobre o espaço de fase, munido da estrutura de parêntese de Poisson, geram uma álgebra de Lie.

8.5 Transformações Canônicas Infinitesimais

Uma transformação canônica infinitesimal é da forma

$$Q_i = q_i + \delta q_i \equiv q_i + \epsilon f_i(q,p,t) \quad , \quad P_i = p_i + \delta p_i \equiv p_i + \epsilon g_i(q,p,t) \ , \tag{8.5.1}$$

onde ϵ é um parâmetro infinitesimal. Substituindo esta equação em (8.2.1) resulta

$$\sum_i \{p_i\, dq_i - (p_i + \epsilon g_i)(dq_i + \epsilon df_i\} = d\Phi \ , \tag{8.5.2}$$

ou, desprezando termos de segunda ordem em ϵ e escrevendo $\Phi = \epsilon F$,

$$\sum_i (g_i\, dq_i + p_i\, df_i) = -dF \ . \tag{8.5.3}$$

Esta equação pode ser reescrita na forma

$$\sum_i (g_i\, dq_i - f_i\, dp_i) = -dG \ , \tag{8.5.4}$$

onde

$$G = \sum_i p_i\, f_i + F \ . \tag{8.5.5}$$

De (8.5.4) deduz-se de imediato

$$f_i = \frac{\partial G}{\partial p_i} \quad , \quad g_i = -\frac{\partial G}{\partial q_i} \ , \tag{8.5.6}$$

ou seja,

$$\delta q_i = \epsilon\, \frac{\partial G}{\partial p_i} \quad , \quad \delta p_i = -\epsilon\, \frac{\partial G}{\partial q_i} \ . \tag{8.5.7}$$

Em termos de parênteses de Poisson estas equações tomam a forma

$$\delta q_i = \epsilon\, \{q_i, G\} \quad , \quad \delta p_i = \epsilon\, \{p_i, G\} \ , \tag{8.5.8}$$

ou, em notação simplética,

$$\delta \mathbf{z} = \epsilon\, \{\mathbf{z}, G\} \ . \tag{8.5.9}$$

A função $G(q, p, t)$ é chamada de função geradora da transformação canônica infinitesimal.

■ **Exercício 8.5.1.** Mostre que (8.5.7) e (8.5.8) são equivalentes. ■

Evolução Temporal Como Transformação Canônica

De fundamental importância é a transformação canônica infinitesimal gerada pela hamiltoniana H com parâmetro $\epsilon = dt$. Neste caso as Eqs.(8.5.8) fornecem

$$\delta q_i = dt\, \{q_i, H\} = \dot{q}_i\, dt \quad , \quad \delta p_i = dt\, \{p_i, H\} = \dot{p}_i\, dt \ , \tag{8.5.10}$$

onde usamos (8.4.4). Conseqüentemente,

$$Q_i(t) = q_i(t) + \delta q_i(t) = q_i(t) + \dot{q}_i(t)\, dt = q_i(t + dt) \ , \tag{8.5.11a}$$

$$P_i(t) = p_i(t) + \delta p_i(t) = p_i(t) + \dot{p}_i(t)\, dt = p_i(t + dt) \ , \tag{8.5.11b}$$

isto é, a transformação canônica infinitesimal gerada por H leva os valores dos qs e ps no instante t em seus valores no instante $t + dt$. Assim, a evolução do sistema mecânico de seu estado no instante t para seu estado no instante $t + dt$ é uma transformação

canônica infinitesimal gerada pela hamiltoniana do sistema. *A evolução temporal num intervalo de tempo finito é uma transformação canônica que consiste numa sucessão de transformações canônicas infinitesimais geradas por* H. Note que, usando (8.5.5) com $G = H$,

$$\Phi = \epsilon F = dt\,(H - \sum_i p_i \dot{q}_i) = -L\,dt \tag{8.5.12}$$

para a transformação canônica infinitesimal gerada por H. Da Definição 8.2.1 vê-se facilmente que a função associada a duas transformações canônicas consecutivas é a soma das funções associadas às transformações individuais. Logo, para a evolução temporal num intervalo de tempo finito,

$$\Phi = -\int_{t_0}^{t} L\,dt = -S \ , \tag{8.5.13}$$

onde S é a ação . Em outras palavras, $-S$ gera uma transformação canônica que evolui o estado do sistema no instante inicial t_0 ao seu estado num instante genérico t. Ainda da Definição 8.2.1, é fácil perceber que $-\Phi$ gera a inversa da transformação canônica gerada por Φ. Logo, sempre que pode ser encarada como função geradora, a ação S engendra uma transformação canônica que traz o estado do sistema no instante genérico t ao seu estado num instante inicial *fixo* t_0, isto é, as novas variáveis canônicas resultantes da transformação gerada por S são constantes. Esta notável propriedade da ação S será confirmada pela teoria de Hamilton-Jacobi, a ser desenvolvida no próximo capítulo.

Variação Infinitesimal de uma Variável Dinâmica Genérica

Retomando a discussão de questões gerais, a variação infinitesimal δu de uma variável dinâmica genérica $u(q,p,t)$ é definida por

$$\delta u = u(q+\delta q, p+\delta p, t) - u(q,p,t) = \sum_i \Big(\frac{\partial u}{\partial q_i}\,\delta q_i + \frac{\partial u}{\partial p_i}\,\delta p_i \Big) \ , \tag{8.5.14}$$

ou, usando (8.5.7),

$$\delta u = \epsilon\,\{u, G\} \ . \tag{8.5.15}$$

Ao escrever (8.5.14) adotamos o *ponto de vista ativo*, ou seja, encaramos a transformação (8.5.1) como levando o ponto (q,p) ao novo ponto $(q+\delta q, p+\delta p)$ do mesmo espaço de fase. Isso induz uma mudança δu na variável dinâmica u de acordo com (8.5.15).

Prova da Identidade de Jacobi

Incidentalmente, esta última equação proporciona uma demonstração elegante e concisa da identidade de Jacobi (Lemos 2000a). Por um lado, dada uma transformação canônica infinitesimal gerada por $G(q,p,t)$, tomando $u = \{E, F\}$ temos

$$\delta\{E,F\} = \epsilon\,\{\{E,F\},G\} \ , \tag{8.5.16}$$

onde $E(q,p,t)$ e $F(q,p,t)$ são duas variáveis dinâmicas quaisquer. Por outro lado, como

o parêntese de Poisson $\{E, F\}$ independe das variáveis canônicas utilizadas para o seu cálculo, sua variação deve-se exclusivamente às variações de E e F, de modo que

$$\delta\{E, F\} = \{\delta E, F\} + \{E, \delta F\} \ , \tag{8.5.17}$$

donde

$$\delta\{E, F\} = \epsilon\{\{E, G\}, F\} + \epsilon\{E, \{F, G\}\} \ . \tag{8.5.18}$$

Comparando (8.5.16) com (8.5.18) e fazendo rearranjos simples, resulta a identidade de Jacobi

$$\{\{E, F\}, G\} + \{\{G, E\}, F\} + \{\{F, G\}, E\} = 0 \ . \tag{8.5.19}$$

Constantes de Movimento e Teorema de Poisson

Tome agora $u = H$ e suponha que G não dependa explicitamente do tempo. Neste caso, a Eq.(8.5.15) torna-se

$$\delta H = \epsilon \{H, G\} = -\epsilon \frac{dG}{dt} \ , \tag{8.5.20}$$

onde usamos (8.4.3). Desta equação inferimos que se H é invariante ($\delta H = 0$) sob uma transformação canônica infinitesimal, sua função geradora é uma constante de movimento. Reciprocamente, cada constante de movimento gera uma transformação canônica infinitesimal que deixa H invariante. Por exemplo, se H não depende de uma determinada variável q_k então $\delta H = 0$ sob a transformação $Q_i = q_i + \epsilon\delta_{ik}$, cuja função geradora é $G = p_k$. Logo, p_k é constante de movimento, resultado já velho conhecido nosso.

Uma propriedade importante das constantes de movimento está contida no resultado a seguir.

Teorema de Poisson. O parêntese de Poisson de duas constantes de movimento é também uma constante de movimento.[2]

Demonstração. Se F e G são constantes de movimento,

$$\{F, H\} + \frac{\partial F}{\partial t} = 0 \quad , \quad \{G, H\} + \frac{\partial G}{\partial t} = 0 \ .$$

Logo,

$$\frac{d}{dt}\{F, G\} = \{\{F, G\}, H\} + \frac{\partial}{\partial t}\{F, G\} = \{\{F, G\}, H\} + \left\{\frac{\partial F}{\partial t}, G\right\} + \left\{F, \frac{\partial G}{\partial t}\right\}$$

$$= \{\{F, G\}, H\} + \{\{H, F\}, G\} + \{\{G, H\}, F\} = 0$$

em virtude da identidade de Jacobi. □

[2]Este teorema e os parênteses de Poisson apareceram numa mesma comunicação publicada por Poisson em 1809, um quarto de século antes do surgimento das equações de Hamilton.

Às vezes é possível gerar novas constantes de movimento pelos parênteses de Poisson de constantes de movimento já conhecidas. Freqüentemente, no entanto, tais parênteses de Poisson são identicamente constantes ou funções das constantes de movimento previamente conhecidas.

■ **Exemplo 8.5.1.** Um projétil move-se no plano xy num campo gravitacional constante com hamiltoniana $H = (p_x^2 + p_y^2)/2m + mgy$. Provar que $F = y - tp_y/m - gt^2/2$ e $G = x - tp_x/m$ são constantes de movimento e obter novas constantes de movimento por intermédio do teorema de Poisson.

Solução. Temos

$$\frac{dF}{dt} = \{F, H\} + \frac{\partial F}{\partial t} = \{y, H\} - \frac{t}{m}\{p_y, H\} - \frac{p_y}{m} - gt = \frac{p_y}{m} - \frac{t}{m}(-mg) - \frac{p_y}{m} - gt = 0$$

e, analogamente,

$$\frac{dG}{dt} = \{G, H\} + \frac{\partial G}{\partial t} = \{x, H\} - \frac{t}{m}\{p_x, H\} - \frac{p_x}{m} = \frac{p_x}{m} - 0 - \frac{p_x}{m} = 0 \ .$$

O parêntese de Poisson $\{F, G\} \equiv 0$ fornece uma constante de movimento trivial. Mas H é constante de movimento, pois não depende explicitamente do tempo, de modo que

$$M = \{F, H\} = \{y, H\} - \frac{t}{m}\{p_y, H\} = \frac{p_y}{m} + gt$$

e também

$$N = \{G, H\} = \{x, H\} - \frac{t}{m}\{p_x, H\} = \frac{p_x}{m}$$

são novas constantes de movimento. Note que as cinco constantes de movimento F, G, M, N, H não podem ser independentes (por quê?). ■

Solução Formal das Equações de Movimento

Se $u(q, p)$ é uma variável dinâmica que não depende explicitamente do tempo, e o mesmo se dá com a hamiltoniana do sistema, é possível obter-se uma solução formal da equação de movimento para u na forma de uma série de potências de t cujos coeficientes envolvem parênteses de Poisson sucessivos entre u e H. Isto realiza de maneira explícita a transformação canônica finita do estado inicial em $t = 0$ ao estado num instante genérico t como uma sucessão de transformações canônicas infinitesimais geradas por H. De fato, temos

$$\left(\frac{du}{dt}\right)_{t=0} = \{u, H\}_{t=0} \equiv \{u, H\}_0 \ , \tag{8.5.21}$$

$$\left(\frac{d^2u}{dt^2}\right)_{t=0} = \left(\frac{d}{dt}\{u, H\}\right)_{t=0} = \{\{u, H\}, H\}_0 \ . \tag{8.5.22}$$

Por indução,

$$\left(\frac{d^n u}{dt^n}\right)_{t=0} = \underbrace{\{\cdots\{\{u,H\},H\},\cdots,H\}}_{n}_0 \ . \tag{8.5.23}$$

Supondo que u admita expansão em série de Taylor em torno de $t=0$, resulta

$$u(t) = u_0 + t\,\{u,H\}_0 + \frac{t^2}{2!}\,\{\{u,H\},H\}_0 + \frac{t^3}{3!}\,\{\{\{u,H\},H\},H\}_0 + \cdots \ , \tag{8.5.24}$$

que é a solução formal da equação de movimento para u. Na maioria das situações é difícil somar esta série de modo a obter uma solução em forma fechada. Em alguns casos simples, todavia, a série pode terminar após o cálculo de um número finito de coeficientes, ou pode ser possível somá-la em termos de funções elementares.

■ **Exemplo 8.5.2.** Dada uma partícula sujeita a uma força constante F, cuja função hamiltoniana é $H = p^2/2m - Fx$, encontrar $x(t)$ usando a Eq.(8.5.24).

Solução. Temos

$$\{x,H\} = \frac{\partial x}{\partial x}\frac{\partial H}{\partial p} - \frac{\partial x}{\partial p}\frac{\partial H}{\partial x} = \frac{p}{m} \ ,$$

$$\{\{x,H\},H\} = \frac{1}{m}\{p,H\} = \frac{1}{m}\left(\frac{\partial p}{\partial x}\frac{\partial H}{\partial p} - \frac{\partial p}{\partial p}\frac{\partial H}{\partial x}\right) = \frac{F}{m} \ .$$

Como $F =$ constante, todos os demais parênteses de Poisson são nulos, de modo que

$$x(t) = x_0 + t\,\{x,H\}_0 + \frac{t^2}{2!}\,\{\{x,H\},H\}_0 = x_0 + \frac{p_0}{m}\,t + \frac{F}{2m}\,t^2 \ ,$$

que é a solução elementar bem conhecida. ■

Vários outros exemplos exatamente solúveis por meio da Eq.(8.5.24) podem ser encontrados em Farina de Souza & Gandelman (1990).

Parênteses de Poisson e Mecânica Quântica

Os parênteses de Poisson são de enorme importância devido ao papel fundamental que desempenham na transição da teoria clássica para a teoria quântica. A chamada *quantização canônica* consiste essencialmente em associar um operador $\hat{A}$ a cada variável dinâmica $A(q,p,t)$, com o comutador de dois operadores correspondendo ao parêntese de Poisson clássico multiplicado por $i\hbar$. Na mecânica quântica na representação de Heisenberg, um operador $\hat{O}_H$ satisfaz

$$\frac{d\hat{O}_H}{dt} = \frac{\partial \hat{O}_H}{\partial t} + \frac{1}{i\hbar}[\hat{O}_H, \hat{H}] \ ,$$

onde $[\hat{A},\hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$ é o comutador de dois operadores. A similaridade entre esta equação e a Eq.(8.4.3) é marcante. A mecânica formulada na linguagem dos parênteses de Poisson é o análogo clássico da teoria quântica na representação de Heisenberg, com o parêntese de Poisson clássico correspondendo ao comutador quântico dividido por $i\hbar$.

Essa correspondência é possível e consistente porque o comutador quântico tem exatamente as mesmas propriedades algébricas que o parêntese de Poisson clássico. A regra de quantização que faz corresponder $i\hbar\{A,B\}$ a $[\hat{A},\hat{B}]$ foi descoberta por Dirac em 1926 (van der Waerden 1967).

8.6 Parênteses de Poisson do Momento Angular

Os parênteses de Poisson das componentes do momento angular $\mathbf{L}$ entre si ou com um vetor arbitrário $\mathbf{F}$ podem ser obtidos sem cálculos diretos por um método extremamente elegante, aparentemente inventado por Julian Schwinger, baseado no papel de $\mathbf{L}$ como gerador de rotações. Uma rotação anti-horária de um ângulo $d\theta$ do vetor posição $\mathbf{r}$ de uma partícula em torno do eixo z é gerada por L_z. Para comprovar isto, note que a referida rotação do vetor (ponto de vista ativo) equivale a uma rotação horária dos eixos coordenados (ponto de vista passivo). Assim, substituindo ϕ por $-d\theta$ na fórmula (3.1.19) de rotação dos eixos coordenados, resulta

$$X = x - yd\theta \quad , \quad Y = xd\theta + y \quad , \quad Z = z \ , \tag{8.6.1}$$

donde

$$\delta x = -yd\theta \quad , \quad \delta y = xd\theta \quad , \quad \delta z = 0 \ . \tag{8.6.2}$$

É imediato que

$$\delta x = d\theta\,\{x, L_z\} \quad , \quad \delta y = d\theta\,\{y, L_z\} \quad , \quad \delta z = d\theta\,\{z, L_z\} \ , \tag{8.6.3}$$

com $L_z = xp_y - yp_x$.

■ **Exercício 8.6.1.** Verifique que (8.6.3) equivale a (8.6.2) e estenda o resultado a um sistema de n partículas, isto é, prove que a transformação canônica infinitesimal $\delta x_i = -y_i d\theta$, $\delta y_i = x_i d\theta$, $\delta z_i = 0$ é gerada pela componente z do momento angular canônico $\mathbf{L} = \sum_k \mathbf{r}_k \times \mathbf{p}_k$. ■

Mais geralmente, o gerador de rotações em torno da direção definida pelo vetor unitário $\hat{\mathbf{n}}$ é $\mathbf{L}\cdot\hat{\mathbf{n}}$ (demonstração: o eixo z sempre pode ser escolhido ao longo do vetor $\hat{\mathbf{n}}$). Portanto, para um vetor $\mathbf{F}$ arbitrário,

$$\delta\mathbf{F} = d\theta\,\{\mathbf{F}, \mathbf{L}\cdot\hat{\mathbf{n}}\} \ . \tag{8.6.4}$$

Por outro lado, sabemos que

$$\delta\mathbf{F} = d\theta\,\hat{\mathbf{n}}\times\mathbf{F} \ , \tag{8.6.5}$$

donde

$$\{\mathbf{F}, \mathbf{L}\cdot\hat{\mathbf{n}}\} = \hat{\mathbf{n}}\times\mathbf{F} \ , \tag{8.6.6}$$

que é o resultado genérico procurado. Ressalte-se que esta última equação vale para qualquer vetor $\mathbf{F}$ *do sistema*, isto é, que possa ser expresso como função unicamente das variáveis canônicas do sistema, porque $\delta\mathbf{F}$ tem seu significado definido pela Eq.(8.5.14). Se $\mathbf{F}$ depender de algum vetor externo – um campo magnético constante, por exemplo – que não é afetado por uma rotação das variáveis canônicas do sistema, a Eq.(8.6.6) não se aplica.

A partir de (8.6.6) diversos resultados interessantes podem ser obtidos. Escolhendo, por exemplo, $\mathbf{F} = \mathbf{p}$ e $\hat{\mathbf{n}} = \hat{\mathbf{k}}$ resulta

$$\{\mathbf{p}, L_z\} = \hat{\mathbf{k}} \times \mathbf{p} = -p_y \hat{\mathbf{i}} + p_x \hat{\mathbf{j}} \ , \tag{8.6.7}$$

ou, em componentes,

$$\{p_x, L_z\} = -p_y \quad , \quad \{p_y, L_z\} = p_x \quad , \quad \{p_z, L_z\} = 0 \ . \tag{8.6.8}$$

Os parênteses de Poisson de $\mathbf{p}$ com L_x e L_y podem ser facilmentes obtidos pelo mesmo procedimento anterior.

■ **Exercício 8.6.2.** (i) Tomando $\mathbf{F} = \mathbf{L}$ e $\hat{\mathbf{n}}$ sucessivamente igual a $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$, prove que

$$\{L_x, L_y\} = L_z \ , \tag{8.6.9}$$

os demais parênteses de Poisson sendo obtidos permutando ciclicamente xyz. (ii) Demonstre o seguinte teorema: se duas componentes do momento angular são conservadas, então a componente restante é necessariamente conservada.[3] ■

Como último resultado interessante, note que um escalar é invariante sob rotações, portanto

$$\{\mathbf{F} \cdot \mathbf{G}, \mathbf{L} \cdot \hat{\mathbf{n}}\} = 0 \tag{8.6.10}$$

quaisquer que sejam os vetores $\mathbf{F}$ e $\mathbf{G}$ do sistema.

■ **Exercício 8.6.3.** Verifique a identidade $\{\mathbf{F} \cdot \mathbf{G}, A\} = \mathbf{F} \cdot \{\mathbf{G}, A\} + \{\mathbf{F}, A\} \cdot \mathbf{G}$ e use-a para demonstrar (8.6.10) diretamente. ■

8.7 Teoremas de Liouville e de Poincaré

Seja $\mathcal{R}$ uma região do espaço de fase (q, p) que é mapeada na região $\mathcal{R}'$ do espaço de fase (Q, P) por uma transformação canônica. O "volume" Ω da região $\mathcal{R}$ é definido por

$$\Omega = \int_{\mathcal{R}} dq_1 \ldots dq_n dp_1 \ldots dp_n \equiv \int_{\mathcal{R}} d^{2n} z \ . \tag{8.7.1}$$

[3] Este teorema só é verdadeiro na ausência de vínculos. Se houver vínculos o teorema não se aplica, como alguns contra-exemplos permitem constatar (Corben & Stehle 1960, p. 229).

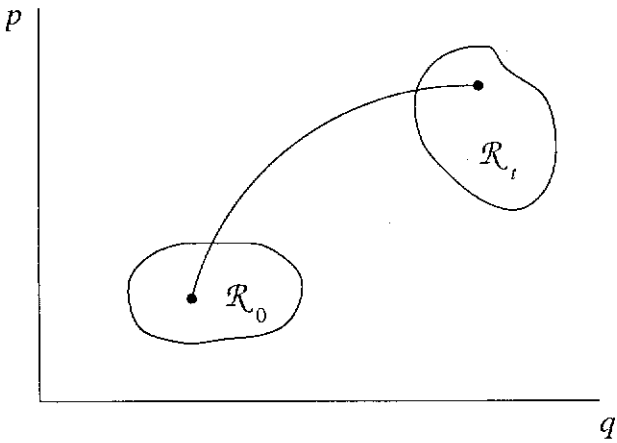


Figura 8.1: Evolução dinâmica de uma região do espaço de fase.

Analogamente,

$$\Omega' = \int_{\mathcal{R}'} dQ_1 \ldots dQ_n dP_1 \ldots dP_n \equiv \int_{\mathcal{R}'} d^{2n}\zeta \ . \tag{8.7.2}$$

Segundo o conhecido teorema de mudança de variáveis numa integral múltipla, temos

$$\Omega' = \int_{\mathcal{R}'} d^{2n}\zeta = \int_{\mathcal{R}} \left| \frac{\partial(\zeta_1, \ldots, \zeta_{2n})}{\partial(z_1, \ldots, z_{2n})} \right| d^{2n}z \ , \tag{8.7.3}$$

o jacobiano da transformação sendo dado por

$$\frac{\partial(\zeta_1, \ldots, \zeta_{2n})}{\partial(z_1, \ldots, z_{2n})} = \begin{vmatrix} \frac{\partial\zeta_1}{\partial z_1} & \cdots & \frac{\partial\zeta_1}{\partial z_{2n}} \\ \vdots & \vdots & \vdots \\ \frac{\partial\zeta_{2n}}{\partial z_1} & \cdots & \frac{\partial\zeta_{2n}}{\partial z_{2n}} \end{vmatrix} = \det \mathbf{M} \ , \tag{8.7.4}$$

onde $\mathbf{M}$ é a matriz jacobiana da transformação canônica $(q, p) \rightarrow (Q, P)$. Mas, de acordo com o Teorema 8.3.2, $|\det \mathbf{M}| = 1$, donde

$$\Omega' = \int_{\mathcal{R}'} d^{2n}\zeta = \int_{\mathcal{R}} d^{2n}z = \Omega \ , \tag{8.7.5}$$

e o volume de qualquer região do espaço de fase é invariante sob transformações canônicas.

Se $\mathcal{R}_0$ é uma região do espaço de fase no instante t_0 com volume Ω_0 e $\mathcal{R}_t$ é a região no instante t obtida levando cada ponto de $\mathcal{R}_0$ no ponto correspondente de $\mathcal{R}_t$ por meio da evolução temporal gerada pela hamiltoniana (Fig. 8.1), então

$$\Omega_t = \Omega_0 \ , \tag{8.7.6}$$

pois a evolução temporal é uma transformação canônica. Esta última equação é conhecida

como teorema de Liouville, o qual assegura que a evolução dinâmica preserva o volume de qualquer região do espaço de fase.

Na mecânica estatística, um ponto do espaço de fase de um sistema clássico (um gás contido num recipiente, por exemplo) é chamado de *ponto representativo* e representa um *microestado* do sistema. Um sistema sob determinadas condições macroscópicas (macroestado) corresponde não a um único estado microscópico, mas a um vasto número de microestados. Uma coleção imaginária de sistemas idênticos em composição e situação macroscópica, mas em diferentes microestados, denomina-se um *ensemble*. Alguns microestados podem ser mais prováveis do que outros, e um ensemble descreve-se por uma densidade $\varrho(q,p,t)$ de pontos representativos no espaço de fase, que serve para caracterizar a distribuição dos diferentes microestados num dado instante. Mais precisamente,

$$\varrho(q,p,t)\, d^n q d^n p \tag{8.7.7}$$

é o número de pontos representativos que no instante t estão contidos no elemento de volume $d^n q d^n p$ do espaço de fase centrado no ponto (q,p). O teorema de unicidade para as equações de Hamilton assegura que o número de pontos representativos num dado elemento de volume é preservado pela evolução temporal. Em outras palavras, cada ponto (q,p) no instante t é produto da evolução dinâmica de um único ponto (q_0,p_0) num instante anterior t_0. Como o volume do elemento também é conservado pela evolução temporal, o teorema de Liouville é equivalente a

$$\frac{d\varrho}{dt} = \frac{\partial \varrho}{\partial t} + \{\varrho, H\} = 0 \ , \tag{8.7.8}$$

e a distribuição de pontos representativos move-se como um fluido incompressível no espaço de fase. Um ensemble é dito estacionário se ϱ não depende explicitamente do tempo, isto é,

$$\frac{\partial \varrho}{\partial t} = 0 \ . \tag{8.7.9}$$

Neste caso, o valor médio de qualquer grandeza física é independente do tempo, de modo que um ensemble estacionário representa um sistema em equilíbrio. Pela Eq.(8.7.8), a densidade ϱ de um sistema em equilíbrio tem parêntese de Poisson zero com H. Esta condição é assegurada tomando-se ϱ como uma função das constantes de movimento do sistema que não dependam explicitamente do tempo. Para sistemas conservativos, uma função qualquer da energia dá conta do recado. O *ensemble microcanônico* é caracterizado por $\varrho(q,p) =$ constante para microestados de uma dada energia e zero em caso contrário, que realiza o postulado das iguais probabilidades *a priori* para um sistema isolado. No caso de um sistema em equilíbrio térmico com um sistema maior, a escolha adequada é $\varrho(q,p) \propto \exp[-H(q,p)/k\Theta]$, que define o *ensemble canônico* (Huang 1963; Pathria 1972).

Embora também desempenhe um papel relevante na concepção de grandes aceleradores de partículas (Corben & Stehle 1960), é na mecânica estatística que o teorema de Liouville encontra suas aplicações mais importantes. Conforme observamos antes, o movimento no espaço de fase assemelha-se ao de um fluido incompressível. Mas um fluido incompressível não pode ter fontes nem sorvedouros. Conseqüentemente, se o espaço de fase disponível para o movimento for limitado, as trajetórias de cada ponto do fluido de fase retornarão quase exatamente sobre si próprias infinitas vezes.

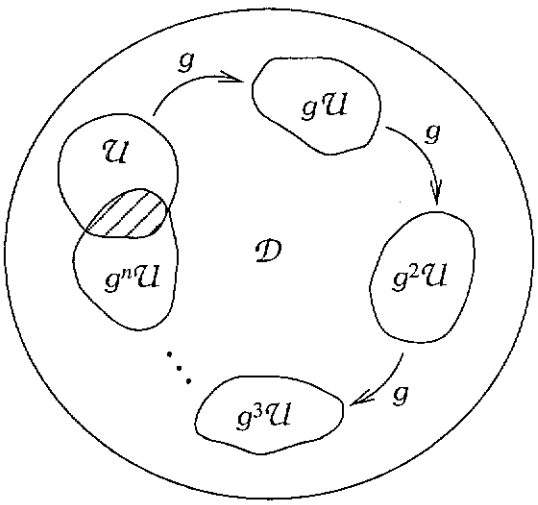


Figura 8.2: Teorema da recorrência de Poincaré.

Teorema da Recorrência de Poincaré. Seja g uma aplicação bijetiva, contínua, que conserva volumes e deixa invariante um domínio limitado $\mathcal{D}$ do espaço de fase de um sistema mecânico: $g\mathcal{D} = \mathcal{D}$. Então em toda vizinhança $\mathcal{U}$ de um ponto qualquer de $\mathcal{D}$ existe um ponto $x \in \mathcal{U}$ que retorna a $\mathcal{U}$ depois de um número finito de aplicações de g, isto é, $g^n x \in \mathcal{U}$ para algum inteiro positivo n.

Demonstração. Como $\mathcal{D}$ é limitado, seu volume é finito. Considere, como na Fig. 8.2, a vizinhança $\mathcal{U}$ e suas imagens:

$$\mathcal{U}\,,\, g\mathcal{U}\,,\, g^2\mathcal{U}\,,\,\ldots\,,\, g^n\mathcal{U}\ .$$

Como todos estes conjuntos possuem o mesmo volume, eles não podem ser disjuntos para todo n, senão o volume de $\mathcal{D}$ seria infinito. Logo, existem números inteiros positivos k, l, com $k > l$, tais que $g^k\mathcal{U} \bigcap g^l\mathcal{U} \neq \emptyset$. Equivalentemente, $g^k x = g^l y$ para algum par de pontos $x, y \in \mathcal{U}$. Aplicando g^{-l} a ambos os lados desta última igualdade resulta $g^{k-l}x = y$ com $x \in \mathcal{U}$ e $y \in \mathcal{U}$. Portanto, $g^n x \in \mathcal{U}$ para $n = k - l > 0$ e a demonstração está completa. □

O teorema é válido, por exemplo, para um gás ideal contido num recipiente de volume V, cujos microestados admissíveis estão restritos ao domínio limitado do espaço de fase definido pelas condições de que todas as N moléculas permaneçam no interior do recipiente e $E \leq H(q,p) \leq E + \Delta E$, ou seja,

$$2mE \leq \mathbf{p}_1^2 + \mathbf{p}_2^2 + \cdots + \mathbf{p}_N^2 \leq 2m(E + \Delta E)\ .$$

Neste caso, a aplicação de interesse é a evolução temporal g^t gerada por $H(q,p)$ que mapeia um ponto do espaço de fase no instante $t = 0$ em sua imagem no instante t, isto é, $g^t\,(q(0),p(0)) = (q(t),p(t))$, e, pelo teorema de Liouville, preserva volumes. Um gás

ideal volta infinitas vezes a um estado arbitrariamente próximo de qualquer estado inicial escolhido, e essa dinâmica quase periódica dá lugar a um aparente paradoxo. Se um gás ideal ocupa a metade de um recipiente e é retirada a parede que impedia a sua expansão para a outra metade, ele rapidamente ocupa todo o volume disponível. De acordo com o teorema da recorrência de Poincaré, depois de "algum tempo" todas as moléculas do gás retornarão à primeira metade do recipiente, em aparente contradição com a segunda lei da termodinâmica.

Dificuldades desse gênero estavam na raiz da controvérsia em que se envolveram, no crepúsculo do século XIX, o matemático Ernst Zermelo[4] e o físico Ludwig Boltzmann, que havia demonstrado a segunda lei da termodinâmica na forma do seu célebre teorema H (Cercignani 1998). Segundo Zermelo, por causa do teorema da recorrência de Poincaré um estado de equilíbrio jamais seria atingido por um gás. A fim de provar que o teorema de Poincaré não tinha as implicações físicas propaladas por Zermelo, Boltzmann estimou o tempo de recorrência para um gás de 10^{18} partículas numa caixa de 1cm^3 como $10^{10^{18}}$ segundos, um número inconcebivelmente grande (a idade atualmente presumida do Universo é "apenas" 10^{17} segundos). Boltzmann argumentou que não havia paradoxo destacando a natureza estatística do seu teorema: os microestados que depois de um certo tempo conduzem a um macroestado de equilíbrio constituem a esmagadora maioria.

Vale sublinhar, também, que nem sempre a conservação da energia é suficiente para definir um domínio de volume finito do espaço de fase que permita a aplicação do teorema da recorrência de Poincaré. No caso em que a energia potencial de interação entre as moléculas do gás pode tornar-se arbitrariamente grande, não são conhecidos domínios invariantes do espaço de fase com volume finito. As trajetórias que escapam para o infinito preenchem uma larga porção do espaço de fase e a elas não se aplica o teorema da recorrência. No entanto, como o fluido de fase não pode ser comprimido, é possível provar que quase todas as trajetórias que vêm do infinito retornam para lá (Thirring 1997).

8.8 Sistemas Hamiltonianos Vinculados

A transição da formulação lagrangiana para a hamiltoniana requer que as velocidades generalizadas possam ser expressas univocamente em termos dos momentos canônicos. Como vimos na Seção 7.1, isso não é possível se a matriz hessiana (7.1.3) é singular. Todas as teorias físicas modernas de significado fundamental são descritas por lagrangianas com matriz hessiana singular, o que justifica uma introdução à dinâmica hamiltoniana de tais sistemas. Nosso tratamento segue de perto a lúcida exposição de Dirac, que vem a ser o criador da teoria dos sistemas hamiltonianos com vínculos (Dirac 1964).

Vínculos Primários

Considere um sistema descrito por uma lagrangiana $L(q, \dot{q})$ sem dependência temporal explícita. Se a matriz hessiana é singular, isto é,

$$\det \mathbf{W} = 0 \ , \tag{8.8.1}$$

então as equações definidoras dos momentos canônicos não podem ser todas resolvidas para as velocidades em função dos momentos porque não constituem um conjunto de

[4] Que se tornaria conhecido pelo *axioma da escolha* e por outras contribuições à teoria dos conjuntos.

equações independentes. Como conseqüência, aparecem relações funcionais entre as coordenadas e momentos do tipo

$$\phi_m(q,p) = 0 \quad , \quad m = 1, \ldots, M \ . \tag{8.8.2}$$

Por causa destas equações, os qs e ps, que são as variáveis dinâmicas básicas do formalismo hamiltoniano, não são mutuamente independentes. As relações (8.8.2), decorrentes unicamente da forma da lagrangiana, são chamadas de *vínculos primários*.

■ **Exemplo 8.8.1.** Dada a lagrangiana

$$L = \frac{1}{2}(\dot{x} - \dot{y})^2 \ , \tag{8.8.3}$$

encontrar os vínculos primários.

Solução. A matriz hessiana

$$\mathbf{W} = \begin{pmatrix} \frac{\partial^2 L}{\partial \dot{x}^2} & \frac{\partial^2 L}{\partial \dot{x} \partial \dot{y}} \\ \frac{\partial^2 L}{\partial \dot{y} \partial \dot{x}} & \frac{\partial^2 L}{\partial \dot{y}^2} \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

é claramente singular. Os momentos canônicos são dados por

$$p_x = \frac{\partial L}{\partial \dot{x}} = \dot{x} - \dot{y} \quad , \quad p_y = \frac{\partial L}{\partial \dot{y}} = \dot{y} - \dot{x} \ . \tag{8.8.4}$$

Estas equações não são independentes entre si, e delas se deduz

$$\phi = p_x + p_y = 0 \ , \tag{8.8.5}$$

que é um vínculo primário. ■

A hamiltoniana é definida da forma usual por $H = \sum_i p_i \dot{q}_i - L$ e a Eq.(7.1.7) permanece válida, mostrando que mesmo na presença de vínculos a hamiltoniana é uma função somente dos qs e ps. Como já sabemos, na presença de vínculos primários devemos empregar a hamiltoniana modificada

$$H_T = H + \sum_{m=1}^{M} \lambda_m \phi_m \ , \tag{8.8.6}$$

que incorpora os vínculos por meio dos multiplicadores de Lagrange λ_m. As equações de Hamilton oriundas do princípio variacional

$$\delta \int_{t_1}^{t_2} \Big(\sum_i p_i \dot{q}_i - H_T \Big) \, dt = \delta \int_{t_1}^{t_2} \Big(\sum_i p_i \dot{q}_i - H - \sum_m \lambda_m \phi_m \Big) \, dt = 0 \ , \tag{8.8.7}$$

com variações arbitrárias e independentes dos qs, ps e λs, são

$$\dot{q}_i = \frac{\partial H}{\partial p_i} + \sum_m \lambda_m \frac{\partial \phi_m}{\partial p_i} \ , \tag{8.8.8a}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} - \sum_m \lambda_m \frac{\partial \phi_m}{\partial q_i} \ , \qquad (8.8.8b)$$

com as condições subsidiárias $\phi_m = 0$ (verifique!).

A fim de escrever as equações de Hamilton (8.8.8) na notação compacta dos parênteses de Poisson, é útil introduzir a noção de *igualdade fraca*, denotada pelo símbolo "$\approx$". Escrevemos inicialmente as equações de vínculo na forma

$$\phi_m(q,p) \approx 0 \qquad (8.8.9)$$

para sublinhar que cada quantidade ϕ_m é restrita a valer zero mas não é identicamente nula como função sobre todo o espaço de fase. Em particular, os ϕ_m têm parênteses de Poisson não nulos com as variáveis canônicas. De modo mais geral, duas funções F, G que coincidem na hipersuperfície do espaço de fase definida pelos vínculos $\phi_m \approx 0$ são ditas *fracamente iguais*, e escrevemos $F \approx G$. Por outro lado, uma igualdade que vale em todo o espaço de fase, e não apenas na hipersuperfície $\phi_m \approx 0$, é dita *forte* e o símbolo usual de igualdade é usado neste caso. No caso do sistema descrito pela lagrangiana (8.8.3), por exemplo, as funções $F(x,y,p_x,p_y) = x+y+(x^2-y^2)\tan(p_x+p_y)$ e $G(x,y,p_x,p_y) = (x+y)e^{p_x+p_y}$ são distintas sobre todo o espaço de fase mas coincidem sobre a hipersuperfície definida pelo vínculo (8.8.5), de modo que $F \approx G$.

Com esses cuidados, a equação de movimento de uma função qualquer $F(q,p)$ pode ser escrita na forma

$$\dot{F} = \{F,H\} + \sum_m \lambda_m \{F,\phi_m\} \approx \{F,H_T\} \ , \qquad (8.8.10)$$

porque cada termo $\{\lambda_m, F\}\phi_m$, sendo proporcional a um dos vínculos, é fracamente zero. A Eq.(8.8.10) mostra que as equações de vínculo (8.8.2) só podem ser usadas *depois* de calculados todos os parênteses de Poisson em que estivermos interessados.

Condições de Consistência e Algoritmo de Dirac-Bergmann

Se tomarmos F como um dos ϕ_m devemos ter $\dot{\phi}_m \approx 0$ por consistência: os vínculos primários devem ser preservados no tempo. Isto dá lugar às *condições de consistência*

$$\{\phi_m, H\} + \sum_{m'} \lambda_{m'} \{\phi_m, \phi_{m'}\} \approx 0 \ . \qquad (8.8.11)$$

Três casos distintos podem ocorrer.

Caso (i). As condições de consistência são identicamente satisfeitas. Neste caso, os únicos vínculos da teoria são os vínculos primários ϕ_m e os multiplicadores de Lagrange são inteiramente arbitrários, de modo que a dinâmica contém funções arbitrárias do tempo.

■ **Exemplo 8.8.2.** Discutir as condições de consistência e obter a solução geral das equações de Hamilton para o sistema descrito pela lagrangiana do Exemplo 8.8.1.

Solução. A hamiltoniana associada à lagrangiana (8.8.3) é

$$H = \dot{x}p_x + \dot{y}p_y - L = (\dot{x}-\dot{y})p_x - \frac{1}{2}p_x^2 = \frac{1}{2}p_x^2 \ , \qquad (8.8.12)$$

onde usamos (8.8.4) e (8.8.5). A condição de consistência

$$\{\phi, H\} + \lambda\{\phi, \phi\} \approx 0 \tag{8.8.13}$$

para o vínculo (8.8.5) é identicamente satisfeita porque $\{\phi, H\} = 0$. Não há outros vínculos e as equações de Hamilton tomam a forma

$$\dot{x} \approx \{x, H\} + \lambda\{x, \phi\} = p_x + \lambda \quad , \quad \dot{p}_x \approx \{p_x, H\} + \lambda\{p_x, \phi\} = 0 \ , \tag{8.8.14a}$$

$$\dot{y} \approx \{y, H\} + \lambda\{y, \phi\} = \lambda \quad , \quad \dot{p}_y \approx \{p_y, H\} + \lambda\{p_y, \phi\} = 0 \ , \tag{8.8.14b}$$

cuja solução geral é $p_x = -p_y = a$, $x(t) = y(t) + at + b$ com a, b constantes arbitrárias e $y(t)$ função arbitrária. ■

■ **Exercício 8.8.1.** Mostre que a solução geral das equações de Lagrange que derivam da lagrangiana (8.8.3) coincide com a solução geral das equações de Hamilton. ■

Caso (ii). As condições de consistência determinam univocamente os multiplicadores de Lagrange. Isto se dá se a matriz $\| \{\phi_m, \phi_{m'}\} \|$ não é singular, isto é, se

$$\det \| \{\phi_m, \phi_{m'}\} \| \not\approx 0 \ . \tag{8.8.15}$$

Neste caso, se $\| C_{mm'} \|$ é a inversa de $\| \{\phi_m, \phi_{m'}\} \|$, temos

$$\sum_{m''} C_{mm''} \{\phi_{m''}, \phi_{m'}\} = \delta_{mm'} \ . \tag{8.8.16}$$

Então as condições de consistência (8.8.11) fornecem

$$\lambda_m \approx -\sum_{m'} C_{mm'} \{\phi_{m'}, H\} \ , \tag{8.8.17}$$

donde

$$\dot{F} \approx \{F, H\} - \sum_{m,m'} \{F, \phi_m\} C_{mm'} \{\phi_{m'}, H\} \ . \tag{8.8.18}$$

Definindo o *parêntese de Dirac*

$$\{F, G\}^* = \{F, G\} - \sum_{m',m''} \{F, \phi_{m'}\} C_{m'm''} \{\phi_{m''}, G\} \ , \tag{8.8.19}$$

a equação de movimento para uma função arbitrária $F(q, p)$ reduz-se a

$$\dot{F} = \{F, H\}^* \ . \tag{8.8.20}$$

Podemos usar o sinal de igualdade forte nesta última equação porque, agora, os parênteses de Dirac dos vínculos com qualquer $F(q, p)$ são nulos:

$$\{F,\phi_m\}^* = \{F,\phi_m\} - \sum_{m',m''} \{F,\phi_{m'}\} C_{m'm''} \{\phi_{m''},\phi_m\} = \{F,\phi_m\} - \sum_{m'} \{F,\phi_{m'}\}\delta_{m'm} = 0 \ . \tag{8.8.21}$$

Segue-se que os vínculos podem ser postos iguais a zero na hamiltoniana H_T antes de calcular os parênteses de Dirac, de modo que $H_T = H$ e as igualdades fracas podem ser tomadas como igualdades fortes. Os parênteses de Dirac têm as mesmas propriedades algébricas que os parênteses de Poisson e satisfazem a identidade de Jacobi.

Caso (iii). As condições de consistência geram vínculos secundários

$$\chi_s(q,p) = 0 \quad , \quad s = 1,\ldots,S \ . \tag{8.8.22}$$

A diferença importante entre vínculos primários e secundários é que os primeiros resultam meramente da forma da lagrangiana e da definição dos momentos, ao passo que os últimos exigem o emprego das equações de movimento. Impomos, agora,

$$\dot{\chi}_s \approx \{\chi_s, H\} + \sum_{m'} \lambda_{m'} \{\chi_s, \phi_{m'}\} \approx 0 \ . \tag{8.8.23}$$

Se recairmos nos casos (i) ou (ii) o processo termina neste estágio. Se isto não acontecer, aparecem outros vínculos secundários e repetimos o procedimento, conhecido como *algoritmo de Dirac-Bergmann*. Após um número finito de estágios o processo termina e ficamos com um conjunto de vínculos secundários denotados por

$$\phi_k(q,p) \approx 0 \quad , \quad k = M+1,\ldots,M+K \ , \tag{8.8.24}$$

onde K é o número total de vínculos secundários. Antecipando que a distinção entre vínculos primários e secundários será pouco importante para a teoria em sua forma final, introduzimos a seguinte notação uniforme:

$$\phi_j(q,p) \approx 0 \quad , \quad j = 1,\ldots,M+K = J \ . \tag{8.8.25}$$

As condições finais de consistência são

$$\{\phi_j, H\} + \sum_{m=1}^{M} \lambda_m \{\phi_j, \phi_m\} \approx 0 \tag{8.8.26}$$

e impõem restrições aos λs, já que não geram novos vínculos. O sistema de J equações lineares não-homogêneas (8.8.26) para as $M \leq J$ incógnitas λ_m deve ser solúvel, caso contrário a dinâmica descrita pela lagrangiana original seria inconsistente, hipótese que excluímos. A solução geral de (8.8.26) é da forma

$$\lambda_m = U_m + \sum_{a=1}^{A} v_a V_m^{(a)} \ , \tag{8.8.27}$$

onde U_m é uma solução particular das equações não-homogêneas[5] e $V_m^{(a)}, a = 1, \ldots, A$, são as soluções linearmente independentes das equações homogêneas, isto é,

$$\sum_{m=1}^{M} \{\phi_j, \phi_m\} V_m^{(a)} \approx 0 \ . \tag{8.8.28}$$

Os coeficientes v_a em (8.8.27) são inteiramente arbitrários.

Substituindo (8.8.27) em (8.8.6) somos conduzidos à hamiltoniana final da teoria que, com Dirac, chamaremos de *hamiltoniana total*:

$$H_T = H + \sum_m U_m \phi_m + \sum_{m,a} v_a V_m^{(a)} \phi_m \equiv H' + \sum_a v_a \Phi_a \ , \tag{8.8.29}$$

onde

$$H' = H + \sum_{m=1}^{M} U_m \phi_m \tag{8.8.30}$$

e definimos

$$\Phi_a = \sum_{m=1}^{M} V_m^{(a)} \phi_m \ . \tag{8.8.31}$$

Em termos da hamiltoniana total as equações de movimento escrevem-se simplesmente

$$\dot{F} \approx \{F, H_T\} \ . \tag{8.8.32}$$

O quadro geral pode agora ser resumido. Os coeficientes v_a são arbitrários, de modo que a solução geral das equações de movimento contém funções arbitrárias do tempo. As condições iniciais não determinam univocamente as variáveis dinâmicas em tempos futuros, mas o estado *físico* do sistema deve ser univocamente determinado pelas condições iniciais. Isto é típico das teorias dotadas de invariância sob transformações de calibre, que em geral são descritas por um número maior de variáveis do que o número de graus de liberdade físicos da teoria.[6] Dado o estado inicial $(q(0), p(0))$ compatível com os vínculos, o estado físico do sistema no instante t é representado pela classe de equivalência dos pontos $(q(t), p(t))$ que só diferem entre si por distintas escolhas dos coeficientes v_a.

Geometricamente, a situação pode ser descrita da seguinte forma (Fig. 8.3). A dinâmica se desenrola na hipersuperfície S do espaço de fase definida pelos vínculos $\phi_j(q,p) \approx 0, j = 1, \ldots, J$. De modo geral, os valores iniciais de (q,p) em $t = 0$ não determinam univocamente seus valores no instante t porque diferentes escolhas dos coeficientes v_a dão lugar a diferentes evoluções temporais das variáveis canônicas. Isto reflete o fato de o número de graus de liberdade físicos ser menor do que o número de variáveis utilizadas para descrever a dinâmica do sistema, de modo que os pontos Q, Q', Q'' representam o mesmo estado físico. Uma escolha específica dos coeficientes arbitrários v_a é chamada de *fixação do calibre*, e torna a dinâmica do sistema unívoca. A fixação do calibre pode ser feita impondo vínculos adicionais cujas condições de consistência determinem completamente os coeficientes v_a.

[5] Isto é, $\sum_m \{\phi_j, \phi_m\} U_m \approx -\{\phi_j, H\}$.

[6] A teoria de calibre arquetípica é a eletrodinâmica de Maxwell descrita em termos dos potenciais escalar e vetor.

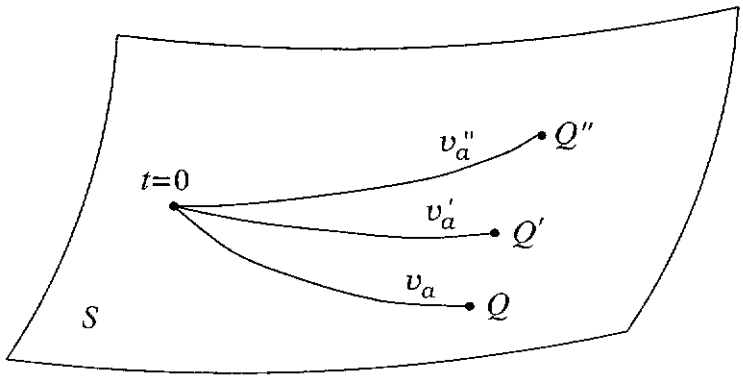


Figura 8.3: Representação geométrica da dinâmica de um sistema hamiltoniano com vínculos.

Vínculos de Primeira Classe e de Segunda Classe

Uma classificação dos vínculos de importância fundamental, especialmente para a transição à teoria quântica, é a que distingue vínculos de *primeira classe* de vínculos de *segunda classe*. Uma função $F(q,p)$ é dita de primeira classe se o seu parêntese de Poisson com qualquer um dos vínculos é fracamente zero:

$$\{F,\phi_j\} \approx 0 \quad , \quad j=1,\ldots,J \ . \tag{8.8.33}$$

Dizemos que F é de segunda classe se o seu parêntese de Poisson com pelo menos um dos vínculos não é fracamente zero. A hamiltoniana total (8.8.29) é de primeira classe porque é a soma de H' de primeira classe com com uma combinação linear com coeficientes arbitrários de vínculos de primeira classe Φ_a definidos pela Eq.(8.8.31). A teoria contém tantas funções arbitrárias do tempo quantos são os vínculos de primeira classe.

■ **Exercício 8.8.2.** Prove que H' e Φ_a, definidas por (8.8.30) e (8.8.31), são funções de primeira classe. ■

Os vínculos primários de primeira classe têm um importante significado físico, que passamos a discutir. Sejam dadas condições iniciais em $t=0$ e consideremos

$$F_{\delta t} = F_0 + \dot{F}\delta t = F_0 + \{F,H_T\}\delta t = F_0 + \delta t\Big[\{F,H'\} + \sum_a v_a\{F,\Phi_a\}\Big] \ . \tag{8.8.34}$$

Como os vs são arbitrários, uma escolha de valores diferentes v_a' para esses coeficientes produz uma mudança em $F_{\delta t}$:

$$\Delta F_{\delta t} = \delta t\sum_a (v_a' - v_a)\{F,\Phi_a\} = \sum_a \epsilon_a\{F,\Phi_a\} \ , \tag{8.8.35}$$

onde

$$\epsilon_a = (v'_a - v_a)\delta t \ . \tag{8.8.36}$$

Como as grandezas físicas devem ser insensíveis a diferentes escolhas dos coeficientes v_a, tanto $F_{\delta t}$ quanto $F_{\delta t} + \Delta F_{\delta t}$ correspondem ao mesmo estado físico. Mas a Eq.(8.8.35) é uma soma de transformações canônicas infinitesimais, cada uma gerada por Φ_a com parâmetro infinitesimal associado ϵ_a. Logo, *os vínculos primários de primeira classe geram transformações de calibre*, isto é, *transformações canônicas infinitesimais que mudam os qs e ps mas não alteram o estado físico do sistema.*

■ **Exemplo 8.8.3.** (Sundermeyer 1982) Aplicar o algoritmo de Dirac-Bergmann à lagrangiana

$$L(q_1, q_2, \dot{q}_1, \dot{q}_2) = \frac{1}{2}\dot{q}_1^2 + q_2\dot{q}_1 + (1-\alpha)q_1\dot{q}_2 + \frac{\beta}{2}(q_1 - q_2)^2 \ . \tag{8.8.37}$$

Solução. A matriz hessiana

$$\mathbf{W} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{8.8.38}$$

é singular. Os momentos canônicos são

$$p_1 = \frac{\partial L}{\partial \dot{q}_1} = \dot{q}_1 + q_2 \quad , \quad p_2 = \frac{\partial L}{\partial \dot{q}_2} = (1-\alpha)q_1 \ , \tag{8.8.39}$$

e dão lugar ao vínculo primário

$$\phi_1 \equiv p_2 + (\alpha - 1)q_1 \approx 0 \ . \tag{8.8.40}$$

A hamiltoniana é dada por

$$H = \frac{1}{2}(p_1 - q_2)^2 - \frac{\beta}{2}(q_1 - q_2)^2 \ . \tag{8.8.41}$$

A exigência de que o vínculo primário seja preservado ao longo do tempo traduz-se na condição de consistência

$$0 \approx \{\phi_1, H\} + \lambda\{\phi_1, \phi_1\} = \alpha(p_1 - q_2) - \beta(q_1 - q_2) \ . \tag{8.8.42}$$

Situação (a): $\alpha = \beta = 0$. A condição de consistência é identicamente satisfeita e não há vínculos secundários. A hamiltoniana total é $H_T = H + \lambda\phi_1$ e as equações de Hamilton são

$$\dot{q}_1 = p_1 - q_2 \quad , \quad \dot{q}_2 = \lambda \quad , \quad \dot{p}_1 = \lambda \quad , \quad \dot{p}_2 = p_1 - q_2 \ . \tag{8.8.43}$$

Como λ é arbitrário, segue-se que q_2 é uma função arbitrária do tempo. Por outro lado, de (8.8.43) deduz-se

$$\ddot{q}_1 = \lambda - \lambda = 0 \implies \dot{q}_1 = \text{constante} \ . \tag{8.8.44}$$

É fácil compreender o porquê desta solução geral das equações de movimento notando que

$$L = \frac{1}{2}\dot{q}_1^2 + \frac{d}{dt}(q_1 q_2) \ . \tag{8.8.45}$$

Uma vez que lagrangianas que só diferem por uma derivada total em relação ao tempo geram as mesmas equações de movimento, somos levados a concluir que o sistema equivale a uma partícula livre com um único grau de liberdade. A variável q_2 não representa nenhum grau de liberdade físico do sistema, é arbitrária e pode ser descartada. A existência de uma função arbitrária do tempo está de acordo com a teoria geral, pois quando só há um vínculo este é necessariamente de primeira classe.

Situação (b): $\alpha = 0, \beta \neq 0$. Neste caso, de (8.8.42) deduz-se um vínculo secundário

$$\phi_2 \equiv q_1 - q_2 \approx 0 \ . \tag{8.8.46}$$

A condição de consistência aplicada a ϕ_2 resulta em

$$\{\phi_2, H\} + \lambda\{\phi_2, \phi_1\} = 0 \implies p_1 - q_2 - \lambda = 0 \implies \lambda = p_1 - q_2 \ , \tag{8.8.47}$$

de modo que o multiplicador de Lagrange fica determinado e a hamiltoniana total torna-se

$$H_T = \frac{1}{2}(p_1 - q_2)^2 - \frac{\beta}{2}(q_1 - q_2)^2 + (p_1 - q_2)(p_2 - q_1) \ . \tag{8.8.48}$$

Os vínculos ϕ_1 e ϕ_2 são de segunda classe porque $\{\phi_1, \phi_2\} = 1$. Um cálculo elementar fornece a matriz dos vínculos e sua inversa:

$$\| \{\phi_a, \phi_b\} \| = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \implies \| C_{ab} \| = \| \{\phi_a, \phi_b\} \|^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \ . \tag{8.8.49}$$

O parêntese de Dirac é dado por

$$\{F, G\}^* = \{F, G\} - \sum_{a,b=1}^{2} \{F, \phi_a\} C_{ab} \{\phi_b, G\} = \{F, G\} + \{F, \phi_1\}\{\phi_2, G\} - \{F, \phi_2\}\{\phi_1, G\} \ . \tag{8.8.50}$$

Com o emprego dos parênteses de Dirac os vínculos podem ser postos iguais a zero na hamiltoniana, que reduz-se a

$$H = H_T = \frac{1}{2}(p_1 - q_2)^2 \ . \tag{8.8.51}$$

As equações de movimento (8.8.20) dão

$$\dot{q}_1 = p_1 - q_2 \quad , \quad \dot{q}_2 = p_1 - q_2 \quad , \quad \dot{p}_1 = p_1 - q_2 \quad , \quad \dot{p}_2 = p_1 - q_2 \ . \tag{8.8.52a}$$

Com o emprego dos parênteses de Dirac os vínculos $p_2 - q_1 = 0$ e $q_2 - q_1 = 0$ tornam-se equações fortes e podem ser substituídas em (8.8.52a) para dar

$$\dot{q}_1 = p_1 - q_1 \quad , \quad \dot{p}_1 = p_1 - q_1 \ . \tag{8.8.52b}$$

A solução geral destas equações é

$$q_1(t) = q_2(t) = p_2(t) = at + b - a \quad , \quad p_1(t) = at + b \ , \tag{8.8.53}$$

com a e b constantes arbitrárias. Confirma-se a análise genérica segundo a qual não restam funções arbitrárias do tempo quando todos os vínculos são de segunda classe.

Se $\alpha \neq 0$, o vínculo secundário derivado de (8.8.42) tem a forma

$$\phi_2 \equiv \alpha(p_1 - q_2) - \beta(q_1 - q_2) \approx 0 \ , \tag{8.8.54}$$

e, com o uso da hamiltoniana (8.8.41), a condição de consistência associada escreve-se

$$\{\phi_2, H\} + \lambda\{\phi_2, \phi_1\} = -\beta\big[(p_1 - q_2) - \alpha(q_1 - q_2)\big] + (\beta - \alpha^2)\lambda \approx 0 \ . \tag{8.8.55}$$

Situação (c): $\alpha \neq 0$, $\beta = \alpha^2$. O vínculo secundário (8.8.54) toma a forma simplificada

$$\phi_2 \equiv p_1 - q_2 - \alpha(q_1 - q_2) \approx 0 \ . \tag{8.8.56}$$

Usando a hamiltoniana (8.8.41) com $\beta = \alpha^2$, a condição de consistência aplicada a ϕ_2 não gera novos vínculos porque

$$\{\phi_2, H\} + \lambda\{\phi_2, \phi_1\} = -\alpha\phi_2 \approx 0 \tag{8.8.57}$$

é identicamente satisfeita em virtude do vínculo $\phi_2 \approx 0$. As equações de movimento (8.8.32) fornecem

$$\dot{q}_1 = p_1 - q_2 + (\alpha - 1)\lambda \ \ , \ \ \dot{q}_2 = \lambda \ , \tag{8.8.58a}$$

$$\dot{p}_1 = \alpha^2(q_1 - q_2) - (\alpha - 1)\lambda \ \ , \ \ \dot{p}_2 = p_1 - q_2 - \alpha^2(q_1 - q_2) \ . \tag{8.8.58b}$$

Resulta que λ (logo, q_2) é função arbitrária do tempo. Isto é consistente com o fato de o vínculo ϕ_1 ser de primeira classe.

Situação (d): $\alpha \neq 0, \beta \neq \alpha^2$. O multiplicador de Lagrange λ é univocamente determinado por

$$\lambda = \frac{\beta}{\beta - \alpha^2}\Big[(p_1 - q_2) - \alpha(q_1 - q_2)\Big] = \frac{\beta}{\alpha}(q_1 - q_2) \ , \tag{8.8.59}$$

onde usamos $\phi_2 \approx 0$. Os vínculos ϕ_1, ϕ_2 são de segunda classe e as equações de movimento são

$$\dot{q}_1 = p_1 - q_2 \ \ , \ \ \dot{q}_2 = \frac{\beta}{\alpha}(q_1 - q_2) \ . \tag{8.8.60a}$$

$$\dot{p}_1 = \frac{\beta}{\alpha}(q_1 - q_2) \ \ , \ \ \dot{p}_2 = \frac{(1-\alpha)\beta}{\alpha}(q_1 - q_2) \ . \tag{8.8.60b}$$

Não aparecem funções arbitrárias do tempo. ■

■ **Exercício 8.8.3.** Na situação (a) do Exemplo 8.8.3, que tipo de transformação canônica infinitesimal é gerada pelo vínculo primário de primeira classe ϕ_1 ? Comprove que essa transformação só afeta quantidades arbitrárias, de modo que trata-se de uma transformação de calibre que não altera o estado físico do sistema. Na situação (d) do Exemplo 8.8.3, determine a forma explícita dos parênteses de Dirac em termos dos parênteses de Poisson e obtenha as equações de movimento (8.8.60). ■

Os exemplos discutidos acima são puramente acadêmicos, mas ilustram bastante bem as diversas situações encontradas nos modelos fisicamente relevantes, os quais consistem

em teorias de campos que ainda não estamos preparados para abordar.[7] Outros aspectos importantes da teoria dos sistemas hamiltonianos vinculados, tais como o papel dos vículos secundários de primeira classe, a hamiltoniana estendida, a conjectura de Dirac, a fixação do calibre e os graus de liberdade independentes, e as regras de quantização de Dirac não cabem neste tratamento introdutório. Para estas e outras questões o leitor interessado deve consultar obras especializadas (Dirac 1964; Henneaux & Teitelboim 1992).

[7]No Capítulo 10 fazemos uma breve discussão, por meio de exemplos simples, de teorias de campos com vínculos.

PROBLEMAS

8.1. (i) Partindo da equação (8.1.9) na forma $\Phi(q,p,t) = F_1(q,Q(q,p,t),t)$, prove que $\partial\Phi/\partial t = -\sum_i P_i \partial Q_i/\partial t + \partial F_1/\partial t$. (ii) Usando $d\Phi(q,p,t) = d\Phi(q,p,t)|_{t\text{ fixo}} + (\partial\Phi/\partial t)\,dt$ e uma equação análoga para $dQ_i(q,p,t)$, mostre que se (8.2.1) vale então (8.1.8) é satisfeita para t variável, com $K(Q,P,t)$ *definida* por (8.1.11).

8.2. Para que valores de α e β a transformação

$$Q = q^\alpha \cos\beta p \quad , \quad P = q^\alpha \operatorname{sen}\beta p$$

é canônica?

8.3. Um certo sistema mecânico com um grau de liberdade tem hamiltoniana

$$H = \frac{1}{2}\Big[(p-aq)^2 + \omega^2(q+bt)^2\Big] \ ,$$

com ω, a e b constantes. (i) Prove que a transformação $Q = q + bt$, $P = p - aq + b$ é canônica e determine uma função geradora. (ii) Mostre que a hamiltoniana transformada é $K = (P^2 + \omega^2 Q^2)/2$. (iii) Usando a solução bem conhecida das equações de movimento nas novas variáveis, retorne às variáveis originais para obter $q(t)$. (d) Prove que

$$R(q,p,t) = \frac{1}{2}(p - aq + b)^2 + \frac{\omega^2}{2}(q+bt)^2$$

é constante de movimento.

8.4. Uma partícula em movimento unidimensional sujeita a uma força constante F possui hamiltoniana $H = p^2/2m - Fq$. (i) Encontre uma transformação canônica tal que a hamiltoniana transformada seja $K(Q,P) = P$. (ii) Resolva as equações de Hamilton para as novas variáveis canônicas e retorne às variáveis originais para obter $q(t)$.

8.5. A hamiltoniana de um sistema com um grau de liberdade tem a forma

$$H = \frac{1}{2}\left(q^4p^2 + \frac{1}{q^2}\right) \ .$$

(i) Escreva as equações de Hamilton e delas deduza uma equação diferencial de segunda ordem envolvendo apenas q. (ii) Invente uma transformação canônica que reduza H à hamiltoniana de um oscilador harmônico: $K = (P^2 + Q^2)/2$. (iii) Sabendo que a solução geral do problema do oscilador é $Q(t) = A\cos(\omega t + \delta)$, determine $P(t)$. (iv) Retornando às variáveis originais, encontre $q(t)$. Esta função é, portanto, a solução da equação diferencial complicada obtida no item (i).

8.6. Resolva a equação de movimento de um oscilador harmônico unidimensional por meio da fórmula (8.5.24), supondo condições iniciais genéricas.

8.7. Um sistema com um grau de liberdade é descrito pela hamiltoniana

$$H = \frac{[p - g(q)]^2}{2f'(q)^2} - kf(q) \ ,$$

onde k é uma constante, f e g sendo funções dadas. Encontre uma transformação canônica que reduza esta hamiltoniana à de uma partícula de massa unitária sujeita a uma força constante. Resolva as equações de Hamilton transformadas e mostre que o movimento do sistema em termos da variável original q é $q(t) = f^{-1}(a + bt + kt^2/2)$, onde a e b são constantes arbitrárias e f^{-1} é a função inversa de f.

8.8. De acordo com o Teorema 2.3.1, as equações de Lagrange não se alteram pela substituição da lagrangiana $L(q, \dot{q}, t)$ por $\bar{L}(q, \dot{q}, t) = L(q, \dot{q}, t) + df(q, t)/dt$, onde $f(q, t)$ é uma função arbitrária. Prove que esta transformação é canônica e encontre uma função geradora.

8.9. Considere a hamiltoniana

$$H = \frac{p_1^2}{2m} + \frac{(p_2 - kq_1)^2}{2m} \ .$$

(i) Determine as constantes A e B de modo que a transformação

$$Q_1 = Ap_1 \quad , \quad P_1 = p_2 - kq_1 \quad , \quad Q_2 = B(p_1 - kq_2) \quad , \quad P_2 = p_2$$

seja canônica. (ii) Mediante a transformação canônica do item anterior e resolvendo as equações de movimento para as variáveis transformadas, encontre a solução das equações de movimento para $q_1(t)$ e $q_2(t)$.

8.10. A hamiltoniana de um certo sistema unidimensional é $H = \omega^2 p(q + t)^2$, onde ω é uma constante positiva. (i) Prove que a transformação $Q = q + t$, $P = p$ é canônica e encontre uma função geradora. (ii) Usando a transformação canônica do item anterior, resolva as equações de Hamilton para $q(t)$ e $p(t)$.

8.11. Uma partícula de massa m move-se ao longo de uma linha reta sob o potencial $V = mk/x^2$. Dadas as condições iniciais $x(0) = x_0$ e $\dot{x}(0) = 0$, obtenha $x(t)$ por meio da solução em série envolvendo parênteses de Poisson sucessivos para $u = x^2$.

8.12. A chamada "mecânica conforme" (Cadoni, Carta & Mignemi 2000) é definida pela hamiltoniana

$$H(q, p) = \frac{p^2}{2f(qp)^2} \ ,$$

onde $f : \boldsymbol{R} \to \boldsymbol{R}$ é uma função diferenciável que nunca se anula. (i) Prove que a transformação

$$Q = qf(qp) \quad , \quad P = \frac{p}{f(qp)}$$

é canônica e, com a sua ajuda, prove que a solução geral das equações de Hamilton para a mecânica conforme é

$$q(t) = \frac{\alpha t + \beta}{f(\alpha^2 t + \alpha\beta)} \quad , \quad p(t) = \alpha f(\alpha^2 t + \alpha\beta) \ ,$$

com α e β constantes arbitrárias. (ii) Mostre que com uma escolha adequada da função

f obtém-se a hamiltoniana do Problema 8.11 e compare a solução acima para $q(t)$ com os resultados obtidos nos Problemas 2.23 e 8.11.

8.13. Mostre que a transformação

$$Q = \ln\left(1 + \sqrt{q}\cos p\right) \quad , \quad P = 2\left(1 + \sqrt{q}\cos p\right)\sqrt{q}\,\mathrm{sen}\,p$$

é canônica e que uma função geradora é $F_3(p, Q) = -\left(e^Q - 1\right)^2 \tan p$.

8.14. (i) Encontre a hamiltoniana associada à lagrangiana $L = e^{\lambda t}(m\dot{q}^2/2 - m\omega^2 q^2/2)$, que descreve um oscilador harmônico amortecido (Problema 1.14). (ii) Prove que a transformação $Q = qe^{\lambda t/2}$, $P = pe^{-\lambda t/2} + (m\lambda/2)qe^{\lambda t/2}$ é canônica e obtenha uma função geradora. (iii) Prove que a transformação

$$Q = \frac{\bar{P}}{m\Omega}\,\mathrm{sen}\left(\frac{m\Omega\bar{Q}}{\bar{P}}\right) \quad , \quad P = \bar{P}\cos\left(\frac{m\Omega\bar{Q}}{\bar{P}}\right)$$

é canônica e reduz a hamiltoniana de um oscilador harmônico de freqüência angular Ω nas variáveis Q, P à hamiltoniana de uma partícula livre nas variáveis $\bar{Q}, \bar{P}$. (iv) Tomando $\Omega = (\omega^2 - \lambda^2/4)^{1/2}$, aplique sucessivamente as tranformações canônicas acima para resolver, no caso de amortecimento fraco, a equação de movimento do oscilador amortecido descrito pela hamiltoniana do item (i). (v) Que modificações são necessárias para tratar os casos de amortecimento forte ($\lambda > 2\omega$) e amortecimento crítico ($\lambda = 2\omega$)?

8.15. A hamiltoniana de um certo sistema é $H = q + te^p$. (i) Mostre que a transformação $Q = q + te^p$, $P = p$ é canônica. (ii) Encontre uma função geradora. (iii) Determine a hamiltoniana transformada $K(Q, P, t)$.

8.16. Um oscilador harmônico bidimensional isotrópico tem hamiltoniana

$$H = \frac{1}{2m}\left(p_x^2 + p_y^2\right) + \frac{m\omega^2}{2}\left(x^2 + y^2\right) \ .$$

Sabendo que $A = p_x^2 + m^2\omega^2 x^2$ e $L = xp_y - yp_x$ são constantes de movimento, use o teorema de Poisson para provar que $B = p_x p_y + m^2\omega^2 xy$ e $C = p_x^2 - p_y^2 + m^2\omega^2(x^2 - y^2)$ também são constantes de movimento.

8.17. Um sistema com dois graus de liberdade é descrito pela hamiltoniana

$$H = p_1^2 + p_2^2 + \frac{1}{2}\left(q_1 - q_2\right)^2 + \frac{1}{8}\left(q_1 + q_2\right)^2 \ .$$

Mostre que a transformação

$$q_1 = \sqrt{Q_1}\cos P_1 + \sqrt{2Q_2}\cos P_2 \quad , \quad q_2 = -\sqrt{Q_1}\cos P_1 + \sqrt{2Q_2}\cos P_2 \ ,$$

$$p_1 = \sqrt{Q_1}\,\mathrm{sen}\,P_1 + \sqrt{Q_2/2}\,\mathrm{sen}\,P_2 \quad , \quad p_2 = -\sqrt{Q_1}\,\mathrm{sen}\,P_1 + \sqrt{Q_2/2}\,\mathrm{sen}\,P_2$$

é canônica e obtenha a hamiltoniana transformada em termos das novas variáveis canônicas. Use os resultados para encontrar q_1, q_2, p_1, p_2 como funções do tempo.

8.18. Ao contrário do que possa parecer, a especificação da função $\Phi(q,p,t)$ na Eq.(8.2.1) não determina univocamente a transformação canônica. No caso de um sistema com apenas um grau de liberdade, mostre que todas as transformações canônicas correspondentes a $\Phi(q,p,t) \equiv 0$ são $Q = f(q)$, $P = p/f'(q)$. As transformações canônicas com $\Phi = 0$ são conhecidas como transformações de Mathieu (Whittaker 1944).

8.19. Este problema consiste numa demonstração alternativa do teorema de Liouville. (i) Seja $\boldsymbol{\mathcal{B}}$ uma matriz tal que $\boldsymbol{\mathcal{B}} = \boldsymbol{I} + \epsilon\boldsymbol{\mathcal{A}}$. Mostre que

$$\det \boldsymbol{\mathcal{B}} = 1 + \epsilon \,\mathrm{tr}\boldsymbol{\mathcal{A}} + \mathcal{O}(\epsilon^2) \ ,$$

onde $\mathcal{O}(\epsilon^2)$ representa termos de ordem superior à primeira em ϵ. (ii) Considere o sistema de equações diferenciais ordinárias

$$\frac{dx_k}{dt} = X_k(x(t),t) \quad , \quad k = 1, \ldots, N$$

e seja $\mathcal{R}_t$ uma região do espaço de fase $(x_1, \ldots, x_N)$ no instante t com volume

$$\Omega(t) = \int_{\mathcal{R}_t} dx_1(t) \cdots dx_N(t) \ .$$

Prove que

$$\frac{d\Omega}{dt} = \int_{\mathcal{R}_t} \boldsymbol{\nabla} \cdot \mathbf{X}\, dx_1(t) \cdots dx_N(t)$$

onde

$$\boldsymbol{\nabla} \cdot \mathbf{X} = \sum_{k=1}^{N} \frac{\partial X_k}{\partial x_k} \ .$$

Sugestão: Mostre que, até primeira ordem em ϵ,

$$x_k(t+\epsilon) = x_k(t) + \epsilon X_k(x(t),t)$$

e note que

$$\Omega(t+\epsilon) = \int_{\mathcal{R}_{t+\epsilon}} dx_1(t+\epsilon) \cdots dx_N(t+\epsilon) = \int_{\mathcal{R}_t} \frac{\partial(x_1(t+\epsilon), \ldots, x_N(t+\epsilon))}{\partial(x_1(t), \ldots, x_N(t))} dx_1(t) \cdots dx_N(t) \ .$$

Em seguida, usando o resultado do item (i), conclua que o jacobiano

$$\frac{\partial(x_1(t+\epsilon), \ldots, x_N(t+\epsilon))}{\partial(x_1(t), \ldots, x_N(t))} = 1 + \epsilon \sum_{k=1}^{N} \frac{\partial X_k(x(t),t)}{\partial x_k(t)} + \mathcal{O}(\epsilon^2)$$

e deduza a expressão desejada para $d\Omega/dt$ recorrendo à definição de derivada. (iii) Prove, finalmente, que para qualquer sistema hamiltoniano $\boldsymbol{\nabla} \cdot \mathbf{X} = 0$, o que estabelece o teorema de Liouville.

Capítulo 9

TEORIA DE HAMILTON-JACOBI

It is impossible to explain honestly the beauties of the laws of nature without some deep understanding of mathematics.

Richard Feynman

No capítulo anterior, vários exemplos foram considerados em que as equações de Hamilton para sistemas mecânicos simples puderam ser resolvidas com a ajuda de transformações canônicas. Tais transformações, todavia, foram obtidas por inspeção, não sendo produto de nenhum procedimento genérico e sistemático. A teoria de Hamilton-Jacobi é um método construtivo que permite, em muitos casos favoráveis, produzir uma transformação canônica capaz de simplificar drasticamente as equações de movimento de um sistema descrito por uma dada função hamiltoniana. Como conseqüência, a integração das equações de movimento transformadas torna-se trivial, reduzindo a resolução das equações de Hamilton a um processo simplesmente algébrico.

9.1 A Equação de Hamilton-Jacobi

Dado um sistema mecânico descrito pelas variáveis canônicas (q, p) e pela hamiltoniana $H(q, p, t)$, efetuemos uma transformação canônica por meio de uma função geradora $S(q, P, t)$. Suponhamos que S possa ser escolhida de tal modo que a hamiltoniana transformada seja nula, isto é, $K(Q, P, t) = 0$. Nessas circunstâncias, as equações de Hamilton transformadas são trivialmente solúveis:

$$\dot{Q}_i = \frac{\partial K}{\partial P_i} = 0 \quad \Longrightarrow \quad Q_i = \beta_i \ , \tag{9.1.1a}$$

$$\dot{P}_i = -\frac{\partial K}{\partial Q_i} = 0 \quad \Longrightarrow \quad P_i = \alpha_i \ , \tag{9.1.1b}$$

onde αs e βs são constantes.

De acordo com (8.1.16) temos

$$0 = K = H(q,p,t) + \frac{\partial S}{\partial t} \ , \tag{9.1.2}$$

donde, usando as equações de transformação

$$p_i = \frac{\partial S}{\partial q_i} \ , \tag{9.1.3}$$

decorrentes da Eq.(8.1.15), deduzimos a *equação de Hamilton-Jacobi*

$$H\Big(q_1,\ldots,q_n,\frac{\partial S}{\partial q_1},\ldots,\frac{\partial S}{\partial q_n},t\Big) + \frac{\partial S}{\partial t} = 0 \ . \tag{9.1.4}$$

Esta é uma equação diferencial parcial de primeira ordem nas $n+1$ variáveis independentes $q_1,\ldots,q_n,t$. Felizmente, não necessitamos de sua solução geral, que envolve uma função arbitrária, e que é muito difícil de ser obtida por tratar-se quase sempre de uma equação não-linear. Como veremos por intermédio de numerosos exemplos, pode ser possível encontrar soluções *particulares* da equação de Hamilton-Jacobi contendo $n+1$ constantes arbitrárias $\alpha_1,\ldots,\alpha_{n+1}$. Uma vez que S não está envolvida diretamente na equação, mas apenas suas derivadas, uma das constantes — digamos, α_{n+1} — é meramente aditiva. Em outras palavras, qualquer solução contendo $n+1$ parâmetros é da forma $S+\alpha_{n+1}$. A constante aditiva α_{n+1} pode ser descartada, pois não modifica a transformação gerada por S. A discussão no parágrafo anterior sugere que, encontrada uma solução da equação de Hamilton-Jacobi da forma $S(q_1,\ldots,q_n,\alpha_1,\ldots,\alpha_n,t)$ onde os αs são constantes de integração não-aditivas, fazendo a identificação $\alpha_i = P_i$ a função $S(q,P,t)$ assim construída executa uma transformação canônica que reduz a zero a nova hamiltoniana. Com essa identificação, o movimento do sistema em termos das variáveis canônicas originais é determinado pelas equações

$$\beta_i = \frac{\partial S}{\partial \alpha_i} \quad , \quad i = 1,\ldots,n \ , \tag{9.1.5}$$

que resultam da combinação de (8.1.15) com (9.1.1a). As n equações (9.1.5) podem ser resolvidas para $q_1,\ldots,q_n$ fornecendo

$$q_i = q_i(\alpha,\beta,t) \ . \tag{9.1.6}$$

Este resultado, conjuntamente com a Eq.(9.1.3), permite escrever

$$p_i = p_i(\alpha,\beta,t) \ . \tag{9.1.7}$$

Estas duas últimas equações representam a solução geral das equações de Hamilton originais, envolvendo $2n$ constantes de integração cujos valores são determinados pelas condições iniciais.

Em face do caráter heurístico da discussão acima, vale a pena torná-la mais rigorosa por meio de definições precisas e demonstrações dos resultados sugeridos anteriormente.

Definição 9.1.1. Uma *solução completa* ou *integral completa* da equação de Hamilton-Jacobi é uma solução da forma $S(q_1,\ldots,q_n,\alpha_1,\ldots,\alpha_n,t)$ contendo n constantes não aditivas $\alpha_1,\ldots,\alpha_n$ e tal que

$$\det\Big(\frac{\partial^2 S}{\partial q_i \partial \alpha_j}\Big) \neq 0 \ . \tag{9.1.8}$$

A condição (9.1.8) caracteriza os parâmetros $\alpha_1,\ldots,\alpha_n$ como mutuamente independentes e, ao mesmo tempo, garante que as Eqs.(9.1.5) podem ser resolvidas para os qs na forma (9.1.6). Podemos, agora, sintetizar o resultado central da teoria de Hamilton-Jacobi na proposição que se segue.

Teorema de Jacobi. Seja $S(q,\alpha,t)$ uma integral completa da equação de Hamilton-Jacobi. Então os qs e ps determinados pelas Eqs.(9.1.5) e (9.1.3) obedecem às equações de Hamilton

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad , \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \quad , \quad i = 1,\ldots,n \ . \tag{9.1.9}$$

Demonstração. Diferenciando (9.1.5) relativamente ao tempo, resulta

$$0 = \dot{\beta}_i = \sum_{j=1}^{n} \frac{\partial^2 S}{\partial q_j \partial \alpha_i}\dot{q}_j + \frac{\partial^2 S}{\partial t \partial \alpha_i} \ . \tag{9.1.10}$$

Por outro lado, lançando mão da Eq.(9.1.3), podemos escrever a equação de Hamilton-Jacobi (9.1.4) na forma

$$H(q,p(q,\alpha,t),t) + \frac{\partial S}{\partial t} = 0 \ . \tag{9.1.11}$$

Diferenciando esta equação em relação a α_i obtemos

$$\sum_{j=1}^{n} \frac{\partial H}{\partial p_j}\frac{\partial p_j}{\partial \alpha_i} + \frac{\partial^2 S}{\partial \alpha_i \partial t} = \sum_{j=1}^{n} \frac{\partial H}{\partial p_j}\frac{\partial^2 S}{\partial \alpha_i \partial q_j} + \frac{\partial^2 S}{\partial \alpha_i \partial t} = 0 \ , \tag{9.1.12}$$

onde empregamos novamente a Eq.(9.1.3). Levando em conta que a ordem de diferenciação é irrelevante, e substituindo (9.1.12) em (9.1.10), ficamos com

$$\sum_{j=1}^{n} \frac{\partial^2 S}{\partial q_j \partial \alpha_i}\Big(\dot{q}_j - \frac{\partial H}{\partial p_j}\Big) = 0 \ . \tag{9.1.13}$$

Como a matriz cujos elemento são $\partial^2 S/\partial q_j \partial \alpha_i$ é não-singular pela Eq.(9.1.8), o sistema homogêneo de equações lineares (9.1.13) possui somente a solução trivial

$$\dot{q}_j - \frac{\partial H}{\partial p_j} = 0 \quad , \quad j = 1,\ldots,n \ . \tag{9.1.14}$$

Tomando a derivada total em relação ao tempo de (9.1.3) vem

$$\dot{p}_i = \sum_{j=1}^{n} \frac{\partial^2 S}{\partial q_j \partial q_i} \dot{q}_j + \frac{\partial^2 S}{\partial t \partial q_i} \ . \tag{9.1.15}$$

Mas de (9.1.11) deduz-se, diferenciando relativamente a q_i,

$$\frac{\partial H}{\partial q_i} + \sum_{j=1}^{n} \frac{\partial H}{\partial p_j} \frac{\partial^2 S}{\partial q_i \partial q_j} + \frac{\partial^2 S}{\partial q_i \partial t} = 0 \ . \tag{9.1.16}$$

A substituição deste último resultado em (9.1.15) conduz a

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} + \sum_{j=1}^{n} \frac{\partial^2 S}{\partial q_i \partial q_j} \left(\dot{q}_j - \frac{\partial H}{\partial p_j} \right) = -\frac{\partial H}{\partial q_i} \tag{9.1.17}$$

em virtude de (9.1.14), completando a demonstração do teorema. □

9.2 Exemplos Unidimensionais

A discussão de alguns exemplos simples servirá para ilustrar o emprego da técnica de Hamilton-Jacobi, além de motivar certas observações de caráter geral que serão feitas mais adiante.

■ **Exemplo 9.2.1.** Resolver a equação de movimento de uma partícula livre unidimensional pelo método de Hamilton-Jacobi.

Solução. Visto que $H = p^2/2m$, a equação de Hamilton-Jacobi assume a forma

$$\frac{1}{2m} \left(\frac{\partial S}{\partial q} \right)^2 + \frac{\partial S}{\partial t} = 0 \ . \tag{9.2.1}$$

Uma integral completa desta equação pode ser obtida por separação de variáveis na forma de soma:

$$S = W(q) + T(t) \ . \tag{9.2.2}$$

Introduzindo S desta forma em (9.2.1) resulta

$$\frac{1}{2m} \left(\frac{dW}{dq} \right)^2 = -\frac{dT}{dt} \ . \tag{9.2.3}$$

Se fixarmos t e variarmos q o primeiro membro de (9.2.3) deveria variar. No entanto, ele não pode variar porque o segundo membro permanece fixo. Logo, ambos os lados de (9.2.3) são iguais a uma mesma constante positiva, que denotaremos por α. Ficamos, assim, com as duas equações diferenciais ordinárias

$$\frac{dW}{dq} = \sqrt{2m\alpha} \quad , \quad -\frac{dT}{dt} = \alpha \ , \tag{9.2.4}$$

donde

$$S(q, \alpha, t) = \sqrt{2m\alpha}\, q - \alpha t \ , \tag{9.2.5}$$

tendo sido abandonadas constantes de integração meramente aditivas. A solução da equação de movimento para q obtém-se de

$$\beta = \frac{\partial S}{\partial \alpha} = \sqrt{\frac{m}{2\alpha}}\, q - t \ , \tag{9.2.6}$$

e de

$$p = \frac{\partial S}{\partial q} = \sqrt{2m\alpha} = \text{constante} = p_0 \ . \tag{9.2.7}$$

Portanto,

$$q = \sqrt{\frac{2\alpha}{m}}\,\beta + \sqrt{\frac{2\alpha}{m}}\, t \equiv q_0 + \frac{p_0}{m}\, t \ , \tag{9.2.8}$$

que é a conhecida solução do problema da partícula livre. ■

■ **Exemplo 9.2.2.** Resolver a equação de movimento de um oscilador harmônico unidimensional pelo método de Hamilton-Jacobi.

Solução. Agora $H = p^2/2m + m\omega^2 q^2/2$, e a equação de Hamilton-Jacobi escreve-se

$$\frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 + \frac{m\omega^2}{2} q^2 + \frac{\partial S}{\partial t} = 0 \ . \tag{9.2.9}$$

Assim como no caso da partícula livre, a equação (9.2.9) é separável na forma

$$S = W(q) - \alpha t \ , \tag{9.2.10}$$

onde W satisfaz

$$\frac{1}{2m}\left(\frac{dW}{dq}\right)^2 + \frac{m\omega^2}{2} q^2 = \alpha \ . \tag{9.2.11}$$

A constante positiva α coincide com o valor constante da hamiltoniana (igual à energia total, neste caso), pois $dW/dq = \partial S/\partial q = p$. A solução de (9.2.11) é imediata:

$$W = \int \sqrt{2m\alpha - m^2\omega^2 q^2}\, dq \ . \tag{9.2.12}$$

Em conseqüência,

$$S(q, \alpha, t) = \int \sqrt{2m\alpha - m^2\omega^2 q^2}\, dq - \alpha t \ , \tag{9.2.13}$$

e a solução da equação de movimento para q obtém-se de

$$\beta = \frac{\partial S}{\partial \alpha} = m \int \frac{dq}{\sqrt{2m\alpha - m^2\omega^2 q^2}} - t = \frac{1}{\omega} \operatorname{sen}^{-1}\left(\sqrt{\frac{m\omega^2}{2\alpha}}\, q\right) - t \ . \tag{9.2.14}$$

Resolvendo esta equação para q encontra-se, com $\delta \equiv \omega\beta$,

$$q(t) = \sqrt{\frac{2\alpha}{m\omega^2}}\, \operatorname{sen}(\omega t + \delta) \ , \tag{9.2.15}$$

que é a bem conhecida solução do problema do oscilador. ■

Nosso último exemplo unidimensional revela que o método de Hamilton-Jacobi pode ser estendido a certas classes de sistemas não-conservativos.

■ **Exemplo 9.2.3.** (Lemos 1979) Usar a teoria de Hamilton-Jacobi para resolver a equação de movimento

$$\ddot{q} + \lambda\dot{q} + \omega^2 q = 0 \ , \tag{9.2.16}$$

que descreve um oscilador harmônico amortecido.

Solução. A Eq.(9.2.16) é gerada pela lagrangiana (vide Problema 1.14)

$$L = e^{\lambda t}\Big(\frac{m\dot{q}^2}{2} - \frac{m\omega^2}{2}q^2\Big) \ , \tag{9.2.17}$$

cuja hamiltoniana correspondente é

$$H = e^{-\lambda t}\frac{p^2}{2m} + \frac{m\omega^2}{2}q^2 e^{\lambda t} \ . \tag{9.2.18}$$

A equação de Hamilton-Jacobi associada a esta hamiltoniana é

$$\frac{e^{-\lambda t}}{2m}\Big(\frac{\partial S}{\partial q}\Big)^2 + \frac{m\omega^2}{2}q^2 e^{\lambda t} + \frac{\partial S}{\partial t} = 0 \ . \tag{9.2.19}$$

Devido à dependência temporal explícita de H, esta última equação não admite separação da variável temporal na forma dos exemplos anteriores (verifique!). No entanto, a forma de H sugere que a transformação

$$Q = qe^{\lambda t/2} \quad , \quad P = pe^{-\lambda t/2} \tag{9.2.20}$$

provavelmente simplificará a hamiltoniana. Acontece que esta transformação é canônica com função geradora $F_2(q,P,t) = e^{\lambda t/2}qP$, e a hamiltoniana transformada

$$K(Q,P,t) = H(q,p,t) + \frac{\partial F_2}{\partial t} = \frac{P^2}{2m} + \frac{m\omega^2}{2}Q^2 + \frac{\lambda}{2}QP \tag{9.2.21}$$

não depende explicitamente do tempo. Assim, o mesmo método dos exemplos anteriores é aplicável ao sistema descrito pelas novas variáveis canônicas (Q,P) e pela hamiltoniana transformada K. A equação de Hamilton-Jacobi associada a K escreve-se

$$\frac{1}{2m}\Big(\frac{\partial S}{\partial Q}\Big)^2 + \frac{m\omega^2}{2}Q^2 + \frac{\lambda}{2}Q\frac{\partial S}{\partial Q} + \frac{\partial S}{\partial t} = 0 \ , \tag{9.2.22}$$

a qual possui solução da forma

$$S = W(Q) - \alpha t \ , \tag{9.2.23}$$

onde

$$\frac{1}{2m}\Big(\frac{dW}{dQ}\Big)^2 + \frac{m\omega^2}{2}Q^2 + \frac{\lambda}{2}Q\frac{dW}{dQ} = \alpha \ . \tag{9.2.24}$$

Note que α é igual ao valor constante da hamiltoniana K. Resolvendo esta equação algébrica do segundo grau para dW/dQ obtém-se

$$\frac{dW}{dx} = -ax \pm \left[b^2 - (1-a^2)x^2\right]^{1/2} \quad . \tag{9.2.25}$$

com $x = (m\omega)^{1/2}Q$, $a = \lambda/2\omega$, $b = (2\alpha/\omega)^{1/2}$. No tocante à resolução de (9.2.16), a escolha do sinal em (9.2.25) é irrelevante (qualquer integral completa serve), de modo que tomaremos

$$W = -\frac{ax^2}{2} + \int \left[b^2 - (1-a^2)x^2\right]^{1/2} dx \quad . \tag{9.2.26}$$

Consideremos o caso $a < 1$, isto é, $\lambda < 2\omega$. Definindo $\gamma = (1-a^2)^{1/2}$ resulta

$$S = -\alpha t - \frac{ax^2}{2} + \int \left(\frac{2\alpha}{\omega} - \gamma^2 x^2\right)^{1/2} dx \quad , \tag{9.2.27}$$

donde

$$\beta = \frac{\partial S}{\partial \alpha} = -t + \frac{1}{\omega}\int \frac{dx}{(b^2 - \gamma^2 x^2)^{1/2}} = -t + \frac{1}{\omega\gamma}\,\mathrm{sen}^{-1}\left(\frac{\gamma x}{b}\right) \quad . \tag{9.2.28}$$

Resolvendo para x, retornando à variável Q e, finalmente, à variável q, resulta

$$q(t) = Ae^{-\lambda t/2}\,\mathrm{sen}(\Omega t + \delta) \quad , \quad \Omega = (\omega^2 - \lambda^2/4)^{1/2} \quad , \tag{9.2.29}$$

onde A e δ são constantes determinadas pelas condições iniciais. A Eq.(9.2.29) é a solução usual para o oscilador amortecido no caso de amortecimento fraco ou subamortecimento. ■

■ **Exercício 9.2.1.** Complete o Exemplo 9.2.3 tratando os casos $\lambda = 2\omega$ (amortecimento crítico) e $\lambda > 2\omega$ (amortecimento forte ou superamortecimento). ■

9.3 Separação de Variáveis

Um aspecto simplificador do método de Hamilton-Jacobi é a desnecessidade de se obter a solução geral da equação de Hamilton-Jacobi para resolver o problema dinâmico, bastando encontrar uma integral completa, que é somente uma solução particular. Soluções particulares de equações diferenciais parciais podem ser comumente obtidas por separação de variáveis. A natureza da equação de Hamilton-Jacobi torna possível, em grande número de casos, a separação de variáveis em forma de soma. Algumas observações de caráter geral, embora extremamente simples, facilitam consideravelmente a tarefa de separar variáveis em boa parte das situações relevantes.

1. Variáveis cíclicas. Se uma certa coordenada, digamos q_n, é variável cíclica de H, é possível separá-la imediatamente das demais. De fato, suponha que H não dependa de q_n. Então podemos escrever

$$S = \alpha_n q_n + \bar{S}(q_1, \ldots, q_{n-1}, t) \tag{9.3.1}$$

onde, substituindo em (9.1.4), $\bar{S}$ satisfaz

$$H\Big(q_1,\ldots,q_{n-1},\frac{\partial\bar{S}}{\partial q_1},\ldots,\frac{\partial\bar{S}}{\partial q_{n-1}},\alpha_n,t\Big)+\frac{\partial\bar{S}}{\partial t}=0\ , \tag{9.3.2}$$

que só envolve as variáveis $q_1,\ldots,q_{n-1},t$. A motivação para buscar uma solução para a equação de Hamilton-Jacobi da forma (9.3.1) provém de

$$\frac{\partial S}{\partial q_n}=p_n=\alpha_n\ , \tag{9.3.3}$$

pois p_n é constante de movimento sempre que q_n é variável cíclica. A solução desta última equação diferencial para S é precisamente da forma (9.3.1). Uma justificativa adicional é que estamos em busca de uma transformação canônica que torne todos os Ps constantes de movimento. Mas p_n já é constante de movimento, de modo que, no que concerne ao par (q_n,p_n), basta executar a transformação identidade, cuja função geradora é $F_2=q_nP_n\equiv q_n\alpha_n$ em virtude da identificação $\alpha_i\equiv P_i$. Se houver mais variáveis cíclicas, é evidente que o primeiro termo à direita do sinal de igualdade em (9.3.1) torna-se uma soma sobre todas as variáveis cíclicas, ao passo que $\bar{S}$ envolve somente as variáveis que não são cíclicas. Por exemplo, se q_n e q_{n-1} são cíclicas temos

$$S=\alpha_nq_n+\alpha_{n-1}q_{n-1}+\bar{S}(q_1,\ldots,q_{n-2},t) \tag{9.3.4}$$

onde

$$H\Big(q_1,\ldots,q_{n-2},\frac{\partial\bar{S}}{\partial q_1},\ldots,\frac{\partial\bar{S}}{\partial q_{n-2}},\alpha_{n-1},\alpha_n,t\Big)+\frac{\partial\bar{S}}{\partial t}=0\ . \tag{9.3.5}$$

2. Variável temporal. Como os exemplos da Seção 9.2 indicam, se H não depende explicitamente do tempo a equação de Hamilton-Jacobi

$$H\Big(q_1,\ldots,q_n,\frac{\partial S}{\partial q_1},\ldots,\frac{\partial S}{\partial q_n}\Big)+\frac{\partial S}{\partial t}=0 \tag{9.3.6}$$

admite separação da variável t na forma

$$S=W(q_1,\ldots,q_n)-\alpha_1t\ , \tag{9.3.7}$$

onde W – chamada de *função característica de Hamilton* – satisfaz

$$H\Big(q_1,\ldots,q_n,\frac{\partial W}{\partial q_1},\ldots,\frac{\partial W}{\partial q_n}\Big)=\alpha_1\ . \tag{9.3.8}$$

Esta equação para W não contém o tempo e α_1 é igual ao valor constante da hamiltoniana. Uma solução completa $S(q,\alpha,t)$ de (9.3.6) fica determinada uma vez que se encontre uma integral completa $W(q,\alpha)$ da *equação de Hamilton-Jacobi independente do tempo* (9.3.8). A função $W(q,\alpha)$, por si só, já produz uma transformação canônica que resolve completamente as equações de movimento para os qs.

■ **Exercício 9.3.1.** Suponha que H não dependa explicitamente do tempo e seja $W(q,\alpha)$ uma integral completa de (9.3.8). Encarando W como função geradora de uma transformação canônica com $\alpha_i \equiv P_i$, mostre que as equações de Hamilton transformadas são

$$\begin{aligned} \dot{P}_i = 0 \quad &\Longrightarrow \quad P_i = \alpha_i \quad , \quad i = 1, \ldots, n \ , \\ \dot{Q}_1 = 1 \quad &\Longrightarrow \quad Q_1 = t + \beta_1 \ , \\ \dot{Q}_i = 0 \quad &\Longrightarrow \quad Q_i = \beta_i \quad , \quad i = 2, \ldots, n \ . \end{aligned} \tag{9.3.9}$$

Em outras palavras, conclua que por meio das equações

$$\frac{\partial W}{\partial \alpha_1} = t + \beta_1 \quad , \quad \frac{\partial W}{\partial \alpha_i} = \beta_i \quad , \quad i = 2, \ldots, n \ , \tag{9.3.10}$$

são determinados os $q_i(\alpha, \beta, t)$ que obedecem às equações de Hamilton e contêm $2n$ constantes de integração cujos valores são fixados pelas condições iniciais. Indicação: $K = H = \alpha_1$. ■

A separabilidade da equação de Hamilton-Jacobi depende do problema físico e do conjunto de coordenadas generalizadas escolhido. Os sistemas para os quais a equação de Hamilton-Jacobi é separável em algum sistema de coordenadas e as equações de movimento são resolvidas por quadraturas pertencem à classe dos chamados *sistemas integráveis* (vide Seção 9.6). O problema de três corpos sob mútua atração gravitacional constitui um dos exemplos mais famosos de sistema não-integrável. Não se conhece nenhum critério completamente geral que indique se e em que sistemas de coordenadas a equação de Hamilton-Jacobi é solúvel por separação de variáveis. Para sistemas de coordenadas generalizadas ortogonais, um teorema devido a Stäckel estabelece as condições necessárias e suficientes para a separabilidade (Pars 1965). Aqui, no entanto, vamos cingir-nos a tratar alguns exemplos de interesse físico que ilustram as técnicas de separação de variáveis costumeiramente empregadas.

■ **Exemplo 9.3.1.** Encontrar uma integral completa da equação de Hamilton-Jacobi para uma partícula no potencial $V(r,\theta) = a(r) + b(\theta)/r^2$ onde (r,θ,φ) são coordenadas esféricas, com $a(r)$ e $b(\theta)$ funções conhecidas.

Solução. A hamiltoniana deste problema é dada por [vide Eq.(7.1.15)]

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\varphi^2}{r^2 \operatorname{sen}^2\theta}\right) + a(r) + \frac{b(\theta)}{r^2} \ , \tag{9.3.11}$$

de modo que a equação de Hamilton-Jacobi toma a forma

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial S}{\partial \theta}\right)^2 + \frac{1}{r^2 \operatorname{sen}^2\theta}\left(\frac{\partial S}{\partial \varphi}\right)^2\right] + a(r) + \frac{b(\theta)}{r^2} + \frac{\partial S}{\partial t} = 0 \ . \tag{9.3.12}$$

Visto que H não depende explicitamente do tempo e φ é coordenada cíclica, podemos escrever

$$S = -\alpha_1 t + \alpha_\varphi \varphi + W(r,\theta) \ , \tag{9.3.13}$$

onde

$$\frac{1}{2m}\left[\left(\frac{\partial W}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial W}{\partial \theta}\right)^2 + \frac{\alpha_\varphi^2}{r^2 \operatorname{sen}^2\theta}\right] + a(r) + \frac{b(\theta)}{r^2} = \alpha_1 \ . \tag{9.3.14}$$

Tentando uma solução para esta equação da forma

$$W(r,\theta) = W_1(r) + W_2(\theta) \tag{9.3.15}$$

somos conduzidos a

$$\frac{1}{2m}\left[\left(\frac{dW_1}{dr}\right)^2 + \frac{1}{r^2}\left(\frac{dW_2}{d\theta}\right)^2 + \frac{\alpha_\varphi^2}{r^2\,\mathrm{sen}^2\theta}\right] + a(r) + \frac{b(\theta)}{r^2} = \alpha_1 \ . \tag{9.3.16}$$

Multiplicando esta equação por r^2 ficamos com

$$r^2\left[\frac{1}{2m}\left(\frac{dW_1}{dr}\right)^2 + a(r) - \alpha_1\right] = -\left[\frac{1}{2m}\left(\frac{dW_2}{d\theta}\right)^2 + \frac{\alpha_\varphi^2}{2m\,\mathrm{sen}^2\theta} + b(\theta)\right] = -\frac{\alpha_\theta^2}{2m} \ , \tag{9.3.17}$$

de modo que as variáveis r e θ foram separadas, a constante de separação tendo sido convenientemente denotada por $-\alpha_\theta^2/2m$. A Eq.(9.3.17) equivale às duas equações diferenciais ordinárias

$$\frac{dW_1}{dr} = \sqrt{2m\big[\alpha_1 - a(r)\big] - \frac{\alpha_\theta^2}{r^2}} \ , \tag{9.3.18}$$

$$\frac{dW_2}{d\theta} = \sqrt{\alpha_\theta^2 - 2mb(\theta) - \frac{\alpha_\varphi^2}{\mathrm{sen}^2\theta}} \ . \tag{9.3.19}$$

Integrando estas equações e inserindo os resultados em (9.3.15) e (9.3.13), resulta uma integral completa de (9.3.12) na forma

$$S = -\alpha_1 t + \alpha_\varphi \varphi + \int\left[2m\big[\alpha_1 - a(r)\big] - \frac{\alpha_\theta^2}{r^2}\right]^{1/2} dr + \int\left[\alpha_\theta^2 - 2mb(\theta) - \frac{\alpha_\varphi^2}{\mathrm{sen}^2\theta}\right]^{1/2} d\theta \ . \tag{9.3.20}$$

A solução das equações de movimento é dada por

$$\beta_1 = \frac{\partial S}{\partial \alpha_1} = -t + m\int \frac{dr}{\left[2m\big[\alpha_1 - a(r)\big] - \alpha_\theta^2/r^2\right]^{1/2}} \ , \tag{9.3.21}$$

$$\beta_\theta = \frac{\partial S}{\partial \alpha_\theta} = -\int \frac{\alpha_\theta dr}{r^2\left[2m\big[\alpha_1 - a(r)\big] - \alpha_\theta^2/r^2\right]^{1/2}} + \int \frac{\alpha_\theta d\theta}{\left[\alpha_\theta^2 - 2mb(\theta) - \alpha_\varphi^2/\,\mathrm{sen}^2\theta\right]^{1/2}} \ , \tag{9.3.22}$$

$$\beta_\varphi = \frac{\partial S}{\partial \alpha_\varphi} = \varphi - \int \frac{\alpha_\varphi d\theta}{\mathrm{sen}^2\theta\left[\alpha_\theta^2 - 2mb(\theta) - \alpha_\varphi^2/\,\mathrm{sen}^2\theta\right]^{1/2}} \ . \tag{9.3.23}$$

É impossível não notar a presteza e simplicidade impressionantes com que se reduz a quadraturas a solução deste problema pelo método de Hamilton-Jacobi. A Eq.(9.3.21) fornece $r(t)$ que, após substituição em (9.3.22), determina $\theta(t)$. Tendo encontrado $\theta(t)$, mediante (9.3.23) obtém-se $\varphi(t)$, completando a resolução das equações de movimento. Caso se esteja interessado apenas

na trajetória geométrica descrita pela partícula, basta utilizar as duas últimas equações para exprimir a equação da curva na forma $r = r(\theta), \varphi = \varphi(\theta)$ ou $\theta = \theta(\varphi), r = r(\varphi)$. ■

Embora os potenciais centrais sejam os mais importantes para as aplicações físicas, o potencial $V(r,\theta) = A/r + B/r^2 \operatorname{sen}^2\theta$, A e B constantes, foi proposto (Hartmann 1972) para descrever o movimento de um elétron em moléculas na forma de um anel, tais como a molécula de benzeno. O potencial de Hartmann é exatamente da forma suposta no exemplo que acabamos de considerar.

■ **Exemplo 9.3.2.** Resolver as equações de movimento de um projétil em três dimensões pelo método de Hamilton-Jacobi.

Solução. Com o eixo z orientado verticalmente para cima,

$$H = \frac{1}{2m}\left(p_x^2 + p_y^2 + p_z^2\right) + mgz \tag{9.3.24}$$

e a equação de Hamilton-Jacobi tem a forma

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2\right] + mgz + \frac{\partial S}{\partial t} = 0 \ . \tag{9.3.25}$$

Como x e y são variáveis cíclicas e H não depende explicitamente do tempo,

$$S = -\alpha_1 t + \alpha_x x + \alpha_y y + W(z) \ , \tag{9.3.26}$$

donde

$$\left(\frac{dW}{dz}\right)^2 = 2m(\alpha_1 - mgz) - \alpha_x^2 - \alpha_y^2 \ . \tag{9.3.27}$$

Integrando esta equação, resulta

$$S = -\alpha_1 t + \alpha_x x + \alpha_y y - \frac{1}{3m^2 g}\left[2m(\alpha_1 - mgz) - \alpha_x^2 - \alpha_y^2\right]^{3/2} \ . \tag{9.3.28}$$

O movimento da partícula obtém-se por meio de

$$\beta_1 = \frac{\partial S}{\partial \alpha_1} = -t - \frac{1}{mg}\left[2m(\alpha_1 - mgz) - \alpha_x^2 - \alpha_y^2\right]^{1/2} \ , \tag{9.3.29a}$$

$$\beta_2 = \frac{\partial S}{\partial \alpha_x} = x + \frac{\alpha_x}{m^2 g}\left[2m(\alpha_1 - mgz) - \alpha_x^2 - \alpha_y^2\right]^{1/2} \ , \tag{9.3.29b}$$

$$\beta_3 = \frac{\partial S}{\partial \alpha_y} = y + \frac{\alpha_y}{m^2 g}\left[2m(\alpha_1 - mgz) - \alpha_x^2 - \alpha_y^2\right]^{1/2} \ . \tag{9.3.29c}$$

A resolução destas três últimas equações para x, y, z fornece

$$x = A + \frac{\alpha_x}{m} t \ , \tag{9.3.30a}$$

$$y = B + \frac{\alpha_y}{m} t \ , \tag{9.3.30b}$$

$$z = C + Dt - \frac{gt^2}{2} \ , \tag{9.3.30c}$$

onde as constantes A, B, C, D exprimem-se em termos dos αs e βs. As Eqs.(9.3.30) coincidem com a solução usual deste problema que se obtém por métodos elementares. ■

9.4 A Ação Como Função das Coordenadas

Conforme observamos na Seção 8.5, a ação desempenha o papel de função geradora de uma transformação canônica que leva o estado do sistema no tempo t ao estado num instante anterior fixo t_0. Ora, isso significa que as variáveis canônicas transformadas são constantes iguais aos valores dos qs e ps no instante t_0. Assim, vê-se que a ação executa uma transformação canônica com as mesmas propriedades que a realizada por uma integral completa S da equação de Hamilton-Jacobi. É de se esperar, portanto, que a ação e S difiram no máximo por uma constante – isto explica o uso da mesma notação S para ambas as quantidades. Tomando a derivada total em relação ao tempo de $S(q, \alpha, t)$ vem

$$\frac{dS}{dt} = \sum_i \frac{\partial S}{\partial q_i}\dot{q}_i + \frac{\partial S}{\partial t} = \sum_i p_i\dot{q}_i - H = L \ , \tag{9.4.1}$$

onde usamos (9.1.3) e (9.1.4). Portanto,

$$S = \int L dt \ , \tag{9.4.2}$$

demonstrando que, vistas como funções do tempo, S e a ação diferem apenas por uma constante. Infelizmente, a Eq.(9.4.2) é absolutamente inútil para determinar $S(q, \alpha, t)$ porque para calcular a integral que nela aparece é preciso conhecer a solução $q_i(t)$ das equações de movimento, mas é exatamente com esta finalidade que se busca S. Deve-se a Hamilton a descoberta original de que a ação, encarada como função das coordenadas, satisfaz a Eq.(9.1.4), que, por isso, também é conhecida como *equação diferencial parcial de Hamilton.*

A *função principal de Hamilton*, denominação que se costuma reservar para a ação expressa em termos das coordenadas, constrói-se de acordo com o seguinte procedimento. Seja $q(\tau)$ um segmento de trajetória física conectando uma dada configuração inicial $q_0 = (q_{01}, \ldots, q_{0n})$ em $\tau = 0$ à configuração $q = (q_1, \ldots, q_n)$ no instante $\tau = t$. A ação correspondente pode ser expressa em termos de q e t, tendo q_0 como parâmetro fixo:

$$S(q, q_0, t) = \int_0^t L\Big(q(\tau), \frac{dq(\tau)}{d\tau}, \tau\Big)\, d\tau \ . \tag{9.4.3}$$

Teorema 9.4.1. A função principal de Hamilton $S(q, q_0, t)$ obedece à equação (9.1.4).

Demonstração. Vejamos, inicialmente, como $S(q, q_0, t)$ varia com q para t fixo. Se a configuração final no instante t for alterada de q para $q + \delta q$, a trajetória física será variada, tornando-se $q(\tau) + \delta q(\tau)$, como indicado na Fig. 9.1. Conseqüentemente, conforme a Eq.(2.3.7),

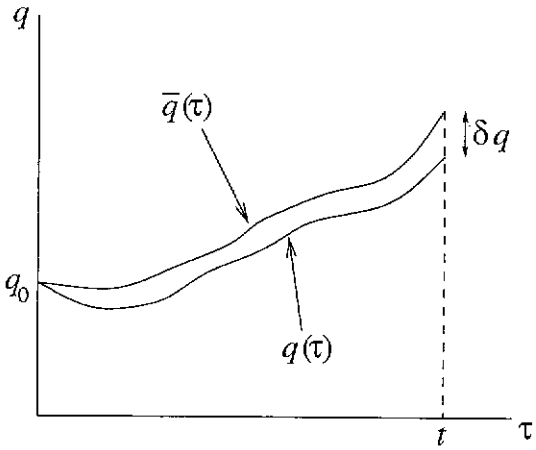


Figura 9.1: Variação envolvida na demonstração do Teorema 9.4.1.

$$\delta S = \int_0^t d\tau \sum_i \left[\frac{\partial L}{\partial q_i} - \frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{q}_i}\right) \right] \delta q_i(\tau) + \sum_i \frac{\partial L}{\partial \dot{q}_i} \delta q_i(\tau) \Big|_{\tau=0}^{\tau=t} = \sum_i p_i \delta q_i \ , \qquad (9.4.4)$$

uma vez que $\delta q_i(\tau = 0) = 0$ e a trajetória física satisfaz as equações de Lagrange.[1] Da Eq.(9.4.4) deduz-se

$$p_i = \frac{\partial S}{\partial q_i} \ . \qquad (9.4.5)$$

Estudemos, agora, como varia S como função de t. Por um lado, de (9.4.3) é claro que

$$\frac{dS}{dt} = L \ . \qquad (9.4.6)$$

Por outro lado,

$$\frac{dS}{dt} = \sum_i \frac{\partial S}{\partial q_i} \dot{q}_i + \frac{\partial S}{\partial t} = \sum_i p_i \dot{q}_i + \frac{\partial S}{\partial t} \ , \qquad (9.4.7)$$

donde

$$\frac{\partial S}{\partial t} = L - \sum_i p_i \dot{q}_i = -H \ . \qquad (9.4.8)$$

Combinando esta equação com (9.4.5) conclui-se que S é solução da Eq.(9.1.4). □

■ **Exemplo 9.4.1.** Construir a função principal de Hamilton para uma partícula sujeita a uma força constante F.

[1] Na Eq.(9.4.4) $\dot{q}_i$ designa $dq_i(\tau)/d\tau$.

Solução. O movimento físico com $x(0) = x_0$ é

$$x(\tau) = x_0 + v_0\tau + \frac{F}{2m}\tau^2 \ . \tag{9.4.9}$$

Impondo a condição $x(t) = x$ obtém-se

$$v_0 = \frac{x - x_0}{t} - \frac{F}{2m}t \ , \tag{9.4.10}$$

donde

$$x(\tau) = x_0 + (x - x_0)\frac{\tau}{t} + \frac{F\tau}{2m}(\tau - t) \ . \tag{9.4.11}$$

Portanto,

$$S(x, x_0, t) = \int_0^t d\tau \left[\frac{m}{2}\left(\frac{dx}{d\tau}\right)^2 + Fx(\tau)\right] = \frac{m(x - x_0)^2}{2t} + \frac{x + x_0}{2}Ft - \frac{F^2t^3}{24m} \ . \tag{9.4.12}$$

Fica ao encargo do leitor verificar que a equação de Hamilton-Jacobi

$$\frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 - Fx + \frac{\partial S}{\partial t} = 0 \tag{9.4.13}$$

é satisfeita pela função principal de Hamilton (9.4.12). ■

Com a identificação $q_{0\,i} \equiv P_i$, a função principal de Hamilton gera uma transformação canônica que reduz as variáveis canônicas transformadas a constantes de movimento. Assim, a determinação de $S(q, q_0, t)$ equivale à resolução completa das equações da Hamilton para as variáveis canônicas originais. Vale sublinhar que a construção de uma função unívoca $S(q, q_0, t)$ só é possível se existir somente uma trajetória dinâmica passando por q_0 e q. Mas isto só pode ser garantido em geral para $|t - t_0|$ suficientemente pequeno, porque neste caso as órbitas físicas que emanam de q_0 em $t = t_0$ não se cruzam. É por isso que a existência de integrais completas da equação de Hamilton-Jacobi só é assegurada *localmente*. Uma questão muito mais delicada é a da construção de uma integral completa válida *globalmente* (para todo t finito), que só é possível para os chamados *sistemas integráveis* (vide Seção 9.6).

Historicamente, coube a Hamilton perceber que, uma vez expressa em termos dos pontos inicial e final de uma trajetória física, a ação satisfaz a equação (9.1.4). No entanto, no tocante apenas à resolução das equações de movimento, a linha seguida por Hamilton é um círculo vicioso, pois a determinação da função principal de Hamilton requer o conhecimento prévio da solução das equações de movimento. O mérito pelo rompimento desse círculo vicioso pertence a Jacobi, que foi o primeiro a compreender que qualquer integral completa da Eq.(9.1.4) dá a solução do problema dinâmico. Isto não significa, contudo, que as idéias de Hamilton tenham sido infrutíferas. A analogia entre óptica e mecânica por ele desenvolvida revelou-se de extrema importância para a formulação da mecânica ondulatória por Schrödinger (vide Seção 9.9). Nessa analogia, a função $S(q, q_0, t)$ desempenha um papel crucial, bem como na formulação de Feynman da mecânica quântica, na qual as quantidades físicas são expressas por meio das chamadas integrais de trajetória (Feynman & Hibbs 1965).

9.5 Variáveis de Ação e Ângulo

Sistemas periódicos são de grande importância em praticamente todos os ramos da física. Há um método poderoso de obter as freqüências de tais sistemas que se baseia na teoria de Hamilton-Jacobi e que não requer a solução detalhada das equações de movimento. Tal método consiste na introdução de certas *variáveis de ação*, a serem definidas mais adiante. Antes, porém, é necessário especificar a classe de problemas tratáveis pelo método.

Sistemas Multiperiódicos

Definição 9.5.1. Um sistema cuja hamiltoniana não depende explicitamente do tempo é dito *separável* se, para algum conjunto de coordenadas generalizadas $(q_1, \ldots, q_n)$, existe uma integral completa da equação de Hamilton-Jacobi independente do tempo (9.3.8) da forma

$$W(q_1, \ldots, q_n, \alpha_1, \ldots, \alpha_n) = W_1(q_1, \alpha_1, \ldots, \alpha_n) + \cdots + W_n(q_n, \alpha_1, \ldots, \alpha_n) \ . \tag{9.5.1}$$

No caso de um sistema separável, temos

$$p_i = \frac{\partial S}{\partial q_i} = \frac{\partial W_i}{\partial q_i} = f_i(q_i, \alpha_1, \ldots, \alpha_n) \ . \tag{9.5.2}$$

Esta última equação é a projeção sobre o plano (q_i, p_i) do movimento que o sistema realiza no espaço de fase.

Definição 9.5.2. Um sistema separável é dito *multiperiódico* se a projeção do movimento sobre cada plano (q_i, p_i) enquadra-se numa das duas seguintes categorias: (i) a curva $p_i = p_i(q_i, \alpha)$ é fechada, isto é, q_i oscila entre dois limites definidos $q_i = a_i$ e $q_i = b_i$ [Fig. 9.2(a)]; (ii) p_i é uma função periódica de q_i, com período q_{i0}, embora q_i não seja função periódica do tempo [Fig. 9.2(b)].

O caso (i) costuma ser chamada de *libração*, um termo tomado de empréstimo da astronomia. Um exemplo óbvio é o oscilador harmônico unidimensional. O caso (ii) é usualmente referido como *rotação*, pois ocorre tipicamente quando q_i é uma coordenada angular. Por exemplo, para um corpo rígido girando livremente em torno de um eixo fixo, sendo q o ângulo de rotação, a cada volta completa q varia de 2π e o estado do sistema se repete.

■ **Exemplo 9.5.1.** Encontrar as projeções nos planos de fase do movimento de um oscilador harmônico bidimensional.

Solução. Em coordenadas cartesianas

$$H = \frac{1}{2m}(p_x^2 + p_y^2) + \frac{k_x}{2}x^2 + \frac{k_y}{2}y^2 \tag{9.5.3}$$

e a equação de Hamilton-Jacobi (9.3.8) é separável na forma $W(x, y) = W_1(x) + W_2(y)$ onde

$$\frac{1}{2m}\left(\frac{dW_1}{dx}\right)^2 + \frac{k_x}{2}x^2 = \alpha_x \ , \tag{9.5.4}$$

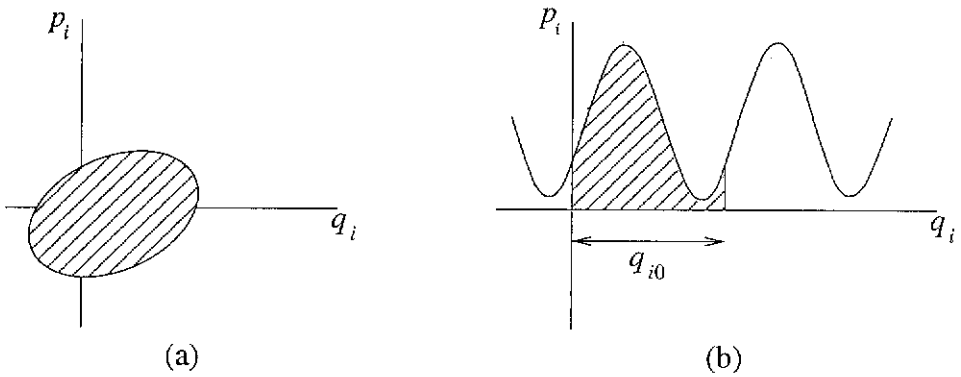


Figura 9.2: Libração (a) ou rotação (b) no plano de fase (q_i, p_i).

$$\frac{1}{2m}\left(\frac{dW_2}{dy}\right)^2 + \frac{k_y}{2}y^2 = \alpha_y \ , \tag{9.5.5}$$

com

$$H = \alpha_1 = \alpha_x + \alpha_y \ , \tag{9.5.6}$$

Como $p_x = dW_1/dx$ e $p_y = dW_2/dy$, as Eqs.(9.5.4) e (9.5.5) representam elipses nos planos de fase (x, p_x) e (y, p_y), respectivamente. Embora ambas as projeções sejam periódicas com períodos $\tau_x = 2\pi/\omega_x = 2\pi(m/k_x)^{1/2}$ e $\tau_y = 2\pi/\omega_y = 2\pi(m/k_y)^{1/2}$, o movimento no espaço de fase só será periódico se as freqüências ω_x e ω_x forem comensuráveis, isto é, se a razão ω_x/ω_y for um número racional. De fato, se τ é o período do movimento, a trajetória no espaço de fase fecha-se no transcurso de um intervalo de tempo τ. As projeções também se fecham, de modo que τ tem que ser um número inteiro de períodos τ_x e τ_y, ou seja, existem inteiros m e n tais que $\tau = m2\pi/\omega_x = n2\pi/\omega_y$, donde $\omega_x/\omega_y = m/n$ é um número racional. ■

O pêndulo simples é um sistema particularmente interessante porque pode exibir libração ou rotação, dependendo do valor da energia. Tomando o nível zero do potencial gravitacional no plano horizontal contendo o ponto de suspensão do pêndulo, temos

$$\frac{p_\theta^2}{2ml^2} - mgl\cos\theta = E \ . \tag{9.5.7}$$

Se $E < mgl$ conclui-se que $-\theta_0 \leq \theta \leq \theta_0$ com $\cos\theta_0 = -E/mgl$. O pêndulo oscila periodicamente entre $-\theta_0$ e θ_0, sendo traçada uma curva fechada no espaço de fase. Se $E > mgl$ qualquer valor de θ é possível, e o pêndulo permanece girando indefinidamente em torno do ponto de suspensão. O movimento se repete cada vez que o ângulo θ varia de 2π, enquadrando-se no caso da Fig. 9.2(b). Se $E = mgl$ o pêndulo encontra-se numa situação limite instável, pois ao atingir a posição vertical $\theta = \pi$ a energia cinética reduz-se a zero e o pêndulo poderia permanecer nessa posição para sempre.[2] Na Fig. 9.3 as curvas 1 e 2 correspondem a $E < mgl$, a curva 3 ao caso limite $E = mgl$ e a curva 4 a $E > mgl$. A curva 3, que separa os casos de libração dos de rotação, é chamada de *separatriz*.

[2]Supondo a massa m suspensa por uma haste rígida de massa desprezível.

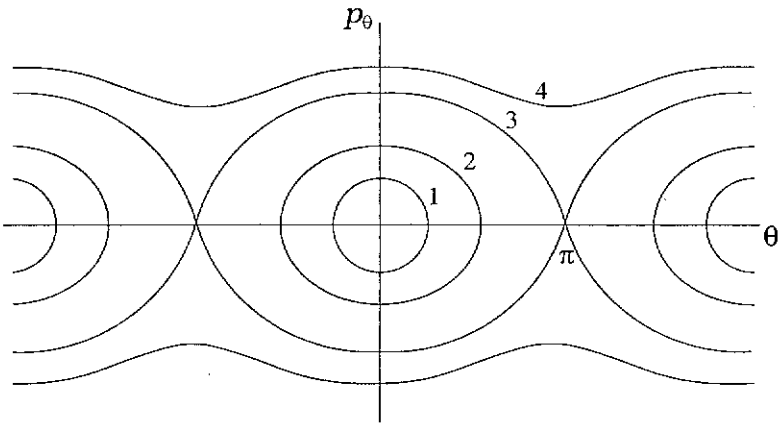


Figura 9.3: Trajetórias no espaço de fase para um pêndulo simples.

Variáveis de Ação-Ângulo

Para sistemas multiperiódicos é possível lançar mão das chamadas *variáveis de ação e ângulo* para efetuar o cálculo das freqüências associadas ao movimento, sem a necessidade de integrar completamente as equações de Hamilton. Esta técnica foi introduzida originalmente por Delaunay na astronomia, onde ela se revela bastante útil porque as freqüências dos vários movimentos podem ser mais interessantes do que o comportamento dinâmico detalhado.

As variáveis de ação para um sistema separável multiperiódico com n graus de liberdade são definidas por

$$J_i = \frac{1}{2\pi} \oint p_i dq_i \quad , \quad i = 1, \ldots, n \ , \tag{9.5.8}$$

onde as integrais estendem-se por um período de libração ou de rotação, conforme o caso. Geometricamente, $2\pi J_i$ representa a área hachurada na Fig. 9.2(a) ou 9.2(b), dependendo do caso. Pela Eq.(9.5.2), os J_i são funções de $\alpha_1, \ldots, \alpha_n$ ou, inversamente, $\alpha_i = \alpha_i(J_1, \ldots, J_n)$. Substituindo estas últimas expressões na integral completa (9.5.1) da equação de Hamilton-Jacobi independente do tempo, resulta uma outra integral completa $W(q_1, \ldots, q_n, J_1, \ldots, J_n)$ onde, agora, as constantes J_i é que são identificadas com os novos momentos P_i. Considerando a transformação canônica gerada por $W(q, J)$, a hamiltoniana transformada é

$$K = H = \alpha_1 = \alpha_1(J_1, \ldots, J_n) \equiv H(J_1, \ldots, J_n) \ , \tag{9.5.9}$$

ou seja, a nova hamiltoniana K é simplesmente a hamiltoniana original H expressa como função das variáveis de ação.

As variáveis canônicas conjugadas aos J_i, conhecidas como *variáveis angulares* φ_i, são definidas por

$$\varphi_i = \frac{\partial W}{\partial J_i} \ . \tag{9.5.10}$$

As equações de movimento para os φ_i escrevem-se

$$\dot{\varphi}_i = \frac{\partial K}{\partial J_i} = \frac{\partial H}{\partial J_i} = \omega_i \ , \tag{9.5.11}$$

onde os ω_i também são constantes, pois só dependem das constantes $J_1, \ldots, J_n$. A solução das equações (9.5.11) é imediata:

$$\varphi_i(t) = \varphi_i(0) + \omega_i t \ . \tag{9.5.12}$$

Freqüências Fundamentais

A fim de elucidar o significado físico dos ωs dados por (9.5.11), suponhamos que o sistema seja periódico com período τ. Evidentemente, a projeção do movimento sobre cada plano de fase também é periódica e as razões das freqüências correspondentes são números racionais. Em outras palavras, depois de um tempo τ cada grau de liberdade terá executado um número inteiro de voltas completas. A mudança correspondente em cada variável angular deve-se à variação das coordenadas q_k, pois os Js são constantes. Portanto, num período do movimento no espaço de fase temos

$$\Delta\varphi_i = \oint \sum_k \frac{\partial \varphi_i}{\partial q_k} dq_k = \oint \sum_k \frac{\partial^2 W}{\partial q_k \partial J_i} dq_k = \frac{\partial}{\partial J_i} \oint \sum_k \frac{\partial W}{\partial q_k} dq_k \ , \tag{9.5.13}$$

onde usamos (9.5.10) e invertemos a ordem de diferenciação. Usando (9.5.1) e (9.5.2), podemos reescrever (9.5.13) na forma

$$\Delta\varphi_i = \frac{\partial}{\partial J_i} \oint \sum_k \frac{\partial W_k}{\partial q_k} dq_k = \frac{\partial}{\partial J_i} \sum_k \oint p_k dq_k \ . \tag{9.5.14}$$

Se n_k é o número de voltas completas descritas pela coordenada q_k num período τ, temos

$$\Delta\varphi_i = \frac{\partial}{\partial J_i} \sum_k n_k \, 2\pi J_k = 2\pi n_i \ . \tag{9.5.15}$$

Cada variável φ_i cresce de um múltiplo inteiro de 2π num período do movimento, justificando considerá-la um ângulo. Por outro lado,

$$\tau = n_i \tau_i \ , \tag{9.5.16}$$

onde τ_i é o período associado ao i-ésimo grau de liberdade, e, pela Eq.(9.5.12),

$$\Delta\varphi_i = \omega_i \tau \ . \tag{9.5.17}$$

Destas três útimas equações deduz-se imediatamente

$$\omega_i \tau_i = 2\pi \quad , \quad i = 1, \ldots, n \ , \tag{9.5.18}$$

de modo que os ω_i são *as freqüências fundamentais do sistema*,[3] isto é, as freqüências do movimento periódico executado por cada grau de liberdade. As derivadas parciais da hamiltoniana em relação às variáveis de ação fornecem as freqüências fundamentais.

■ **Exemplo 9.5.2.** Encontrar as freqüências fundamentais de um oscilador harmônico bidimensional pelo método das variáveis de ação.

Solução. Levando em conta que $dW_1/dx = p_x$, a Eq.(9.5.4) admite duas soluções para p_x, a saber,

$$p_x^{\pm} = \pm\sqrt{2m\alpha_x - mk_x x^2} \ , \tag{9.5.19}$$

na qual x só toma valores entre $x_- = -\sqrt{2\alpha_x/k_x}$ e $x_+ = \sqrt{2\alpha_x/k_x}$. Um ciclo completo da coordenada x consiste numa variação desde x_- até x_+ e de volta a x_-. Na ida de x_- a x_+ a variável x é crescente e $p_x > 0$, com $p_x < 0$ na volta de x_+ a x_-. Portanto,

$$2\pi\, J_x = \oint p_x\, dx = \int_{x_-}^{x_+} p_x^+\, dx + \int_{x_+}^{x_-} p_x^-\, dx = 2\int_{x_-}^{x_+} \sqrt{2m\alpha_x - mk_x x^2}\, dx \ . \tag{9.5.20}$$

Esta integral calcula-se facilmente por meio da substituição $x = \sqrt{2\alpha_x/k_x}\,\mathrm{sen}\,\theta$, dando como resultado

$$J_x = \frac{1}{2\pi} 2 \int_{-\pi/2}^{\pi/2} \sqrt{mk_x}\,\frac{2\alpha_x}{k_x}\cos^2\theta\, d\theta = \frac{1}{\pi}\sqrt{mk_x}\,\frac{2\alpha_x}{k_x}\,\frac{\pi}{2} = \sqrt{\frac{m}{k_x}}\,\alpha_x \ . \tag{9.5.21}$$

O procedimento genérico acima delineado não é necessário no caso do oscilador harmônico. O resultado anterior pode ser obtido praticamente sem cálculos notando que é possível escrever a Eq.(9.5.4) na forma

$$\frac{x^2}{a^2} + \frac{p_x^2}{b^2} = 1 \quad , \quad a = \sqrt{2\alpha_x/k_x} \quad , \quad b = \sqrt{2m\alpha_x} \ , \tag{9.5.22}$$

que é a forma padrão da equação de uma elipse no plano de fase (x, p_x). A variável de ação J_x multiplicada por 2π coincide com a área dessa elipse, isto é,

$$J_x = \frac{1}{2\pi}\pi ab = \sqrt{\frac{m}{k_x}}\,\alpha_x \ . \tag{9.5.23}$$

Analogamente, $J_y = \sqrt{m/k_y}\,\alpha_y$ e, conseqüentemente,

$$H = \alpha_1 = \alpha_x + \alpha_y = \sqrt{\frac{k_x}{m}}\,J_x + \sqrt{\frac{k_y}{m}}\,J_y \ , \tag{9.5.24}$$

donde

$$\omega_x = \frac{\partial H}{\partial J_x} = \sqrt{\frac{k_x}{m}} \quad , \quad \omega_y = \frac{\partial H}{\partial J_y} = \sqrt{\frac{k_y}{m}} \ , \tag{9.5.25}$$

que é o resultado correto para as freqüências das oscilações nas direções x, y. ■

[3] O termo "freqüência" está sendo empregado aqui no sentido de "freqüência angular".

■ **Exemplo 9.5.3.** No caso de um oscilador harmônico unidimensional, obter a transformação canônica das variáveis (x,p) para as variáveis de ação-ângulo (φ, J).

Solução. Por (9.2.12), $W = \int (2m\alpha - m^2\omega^2 x^2)^{1/2}\, dx$. De acordo com o Exemplo 9.5.2, $J = \sqrt{m/k}\,\alpha$ ou $E = \alpha = \omega J$. Assim, $W(x,J) = \int (2m\omega J - m^2\omega^2 x^2)^{1/2}\, dx$, donde

$$\varphi = \frac{\partial W}{\partial J} = \int \frac{m\omega dx}{\sqrt{2m\omega J - m^2\omega^2 x^2}} = \operatorname{sen}^{-1}\left(\sqrt{\frac{m\omega}{2J}}x\right) \quad , \quad p = \frac{\partial W}{\partial x} = \sqrt{2m\omega J - m^2\omega^2 x^2}\ . \tag{9.5.26}$$

Portanto,

$$x = \sqrt{\frac{2J}{m\omega}}\operatorname{sen}\varphi \quad , \quad p = \sqrt{2m\omega J}\cos\varphi \tag{9.5.27}$$

é a transformação canônica procurada. ■

9.6 Sistemas Integráveis e Teorema KAM

A noção de *integrabilidade* de um sistema mecânico refere-se à possibilidade de resolução explícita de suas equações de movimento. Por exemplo, no caso de um sistema conservativo unidimensional constituído por uma só partícula, o fato de a energia ser constante permite escrever a solução da equação de movimento na forma

$$t - t_0 = \sqrt{\frac{m}{2}}\int_{x_0}^{x} \frac{d\xi}{\sqrt{E - V(\xi)}}\ . \tag{9.6.1}$$

Basta agora uma inversão adicional para se obter a solução $x(t)$: o problema de uma partícula em movimento unidimensional sob a ação de uma força conservativa pode ser resolvido por uma quadratura. O pião simétrico com um ponto fixo é um sistema com três graus de liberdade que possui três constantes de movimento independentes: p_ϕ, p_ψ e H. O parêntese de Poisson de qualquer par dessas três constantes de movimento é nulo. Como vimos na Seção 4.10, as equações de movimento do pião resolvem-se por quadraturas.

Definição 9.6.1. As m variáveis dinâmicas $F_1(q,p), \ldots, F_m(q,p)$ estão *em involução* se o parêntese de Poisson de quaisquer duas delas é zero:

$$\{F_k, F_l\} = 0 \quad , \quad k,l = 1,\ldots,m\ . \tag{9.6.2}$$

No Exemplo 9.3.2 do projétil em três dimensões, p_x, p_y e H são constantes de movimento em involução. Uma integral completa da equação de Hamilton-Jacobi associada foi facilmente obtida por separação de variáveis, e as equações de movimento foram resolvidas por quadraturas. De maneira geral, caso, por qualquer método, consigamos encontrar uma integral completa $S(q,\alpha,t) = W(q,\alpha) - \alpha_1 t$ da equação de Hamilton-Jacobi para um sistema conservativo com n graus de liberdade, os momentos canônicos transformados $P_k = \alpha_k(q,p)$ obtidos por inversão de $p_k = \partial W/\partial q_k$ formam um conjunto de n constantes de movimento claramente em involução e as equações de movimento resolvem-se por quadraturas. Parece, portanto, que a integrabilidade de um sistema com n graus de liberdade está intimamente associada à existência de n constantes de movimento em involução. Poisson, em 1837, e Liouville, em 1840, demonstraram que se um

sistema hamiltoniano com dois graus de liberdade possui duas constantes de movimento independentes em involução, então suas equações de movimento podem ser resolvidas por quadraturas. O teorema foi posteriormente generalizado por Liouville, em 1855, e, no século XX, enriquecido pela contribuição de Arnold.

Lema 9.6.1 Se um sistema hamiltoniano conservativo com n graus de liberdade admite n constantes de movimento em involução $F_1(q,p),\ldots,F_n(q,p)$ e a matriz com elementos $W_{kl}=\partial F_k/\partial p_l$ é não-singular, existem variáveis canônicas $(\varphi_1,\ldots,\varphi_n,J_1,\ldots,J_n)$ tais que $H=H(J_1,\ldots,J_n)$.

Demonstração. Aqui nos baseamos em Pars (1965) e Berry (1978). Sejam $F_1,\ldots,F_n$ constantes de movimento em involução com $\det(\partial F_k/\partial p_l)\neq 0$. Defina os candidatos a novos momentos $J_1,\ldots,J_n$ por

$$J_k=F_k(q,p)\quad,\quad k=1,\ldots,n\ . \tag{9.6.3}$$

Conforme o teorema da função implícita, estas equações podem ser resolvidas para os ps na forma

$$p_k=\psi_k(q,J)\quad,\quad k=1,\ldots,n\ . \tag{9.6.4}$$

Portanto,

$$F_k\big(q,\psi(q,J)\big)\equiv J_k\quad,\quad k=1,\ldots,n \tag{9.6.5}$$

donde, diferenciando em relação a q_l, vem

$$\frac{\partial F_k}{\partial q_l}+\sum_s\frac{\partial F_k}{\partial p_s}\frac{\partial\psi_s}{\partial q_l}=0\quad\Longrightarrow\quad\frac{\partial F_k}{\partial q_l}=-\sum_s W_{ks}\frac{\partial\psi_s}{\partial q_l}\ . \tag{9.6.6}$$

Por outro lado, explicitando $\{F_k,F_l\}=0$ temos

$$\sum_j\left(\frac{\partial F_k}{\partial q_j}\frac{\partial F_l}{\partial p_j}-\frac{\partial F_k}{\partial p_j}\frac{\partial F_l}{\partial q_j}\right)=0\ , \tag{9.6.7}$$

que, com o emprego de (9.6.6), torna-se

$$-\sum_{j,s}W_{ks}\frac{\partial\psi_s}{\partial q_j}W_{lj}+\sum_{j,s}W_{kj}\frac{\partial\psi_s}{\partial q_j}W_{ls}=0 \tag{9.6.8}$$

ou, com a troca $j\leftrightarrow s$ dos índices mudos do segundo somatório,

$$\sum_{j,s}W_{ks}\left(\frac{\partial\psi_s}{\partial q_j}-\frac{\partial\psi_j}{\partial q_s}\right)W_{lj}=0\ . \tag{9.6.9}$$

Definindo a matriz $\mathbf{X}$ com elementos

$$X_{rs}=\frac{\partial\psi_r}{\partial q_s}-\frac{\partial\psi_s}{\partial q_r}\ , \tag{9.6.10}$$

podemos exprimir (9.6.9) em notação matricial como

$$\mathbf{W}\mathbf{X}\mathbf{W}^T = 0 \ . \tag{9.6.11}$$

Como $\mathbf{W}$ tem inversa, multiplicando esta equação pela esquerda por $\mathbf{W}^{-1}$ e pela direita por $(\mathbf{W}^T)^{-1}$ resulta $\mathbf{X} = 0$, isto é,

$$\frac{\partial \psi_r}{\partial q_s} = \frac{\partial \psi_s}{\partial q_r} \ . \tag{9.6.12}$$

De acordo com o Apêndice D, existe uma função $\Phi(q, J)$ tal que

$$\psi_r(q, J) = \frac{\partial \Phi(q, J)}{\partial q_r} \ . \tag{9.6.13}$$

Assim, pela Eq.(8.1.15), a transformação $(q, p) \to (\varphi, J)$ definida por

$$\varphi_r = \frac{\partial \Phi}{\partial J_r} \quad , \quad p_r = \frac{\partial \Phi}{\partial q_r} \tag{9.6.14}$$

é canônica com função geradora $\Phi(q, J)$. Como os Js são constantes de movimento, as equações de Hamilton transformadas $\dot{J}_k = -\partial H/\partial \varphi_k$ estabelecem que $\partial H/\partial \varphi_k = 0$. Logo, H não depende dos φs e a demonstração está completa. □

Definição 9.6.2. Um sistema hamiltoniano conservativo com n graus de liberdade e hamiltoniana $H(q, p)$ é dito *completamente integrável* ou, simplesmente, *integrável* se existem n constantes de movimento independentes em involução, isto é:
(a) $\{F_i, H\} = 0 \ , \ i = 1, \ldots, n$;
(b) $\{F_i, F_j\} = 0 \ , \ i, j = 1, \ldots, n$;
(c) os vetores ∇F_i são linearmente independentes[4] em cada ponto do espaço de fase.

Agora estamos prontos para enunciar o resultado central a respeito de sistemas hamiltonianos integráveis.

Teorema de Liouville-Arnold. Considere um sistema hamiltoniano integrável com n graus de liberdade. Então:
(a) Existem variáveis canônicas $(\varphi_1, \ldots, \varphi_n, J_1, \ldots, J_n)$ tais que $H = H(J_1, \ldots, J_n)$, de modo que a solução das equações de movimento nas novas variáveis é

$$J_k = \text{constante} \ , \ \varphi_k = \varphi_k(0) + \omega_k t \ , \ k = 1, \ldots n \ , \tag{9.6.15}$$

onde cada $\omega_k = \partial H/\partial J_k$ é constante.
(b) As equações de Hamilton nas variáveis originais (q, p) podem ser resolvidas por quadraturas.
(c) Se o conjunto das hipersuperfícies de nível das constantes de movimento em involução F_k, definido por

$$\mathcal{M}_C = \left\{ (q, p) \,\middle|\, F_k(q, p) = C_k \ , \ k = 1, \ldots, n \right\} \ , \tag{9.6.16}$$

[4] Aqui $\nabla = (\partial/\partial q_1, \ldots, \partial/\partial q_n, \partial/\partial p_1, \ldots, \partial/\partial p_n)$ é o operador nabla no espaço de fase.

é compacto e conexo,[5] as variáveis canônicas (J, φ) podem ser escolhidas como variáveis de ação-ângulo e o movimento é multiperiódico com freqüências $\omega_k = \partial H/\partial J_k$.

Demonstração. Um resultado da álgebra linear (Dettman 1974, Teorema 2.6.5) e alguns momentos de reflexão mostram que a condição (c) da Definição 9.6.2 equivale a dizer que uma das matrizes quadradas com elementos $\partial F_i/\partial z_s$, onde s toma n valores no conjunto $\{1, 2, \ldots, 2n\}$, é não-singular. Se alguns dos zs coincidirem com alguns dos qs, podemos fazer uma transformação canônica que troca cada um desses qs pelo seu momento conjugado (vide Exercício 8.1.2). Portanto, sem perda de generalidade, podemos supor que $s = n+1, \ldots, 2n$, de modo que $\det(\partial F_i/\partial p_j) \neq 0$. Pelo Lema 9.6.1, existem variáveis canônicas (φ, J), com $J_k = F_k(q,p)$, tais que a hamiltoniana transformada só depende dos Js e as equações de Hamilton nas novas variáveis são solúveis na forma (9.6.15), o que prova (a). Para obter a função geradora $\Phi(q, J)$ da transformação canônica $(q,p) \to (\varphi, J)$ basta resolver o sistema de equações (9.6.3) para os ps e integrar ao longo de uma linha reta uma função construída com as funções supostamente conhecidas $\psi_k(q, J)$ (vide Eq.(D.7) do Apêndice D). Assim, a determinação das variáveis canônicas originais como funções do tempo só requer a inversão e a integração de funções conhecidas, de modo que as equações de Hamilton originais resolvem-se por quadraturas. Isto prova (b). A parte (c) refere-se à estrutura global do espaço de fase e sua demonstração será omitida, pois requer o recurso a argumentos topológicos, de teoria dos grupos e de geometria diferencial com os quais não supomos que o leitor esteja familiarizado. O leitor interessado numa prova detalhada de (c) deve consultar a literatura pertinente (Arnold 1976; Thirring 1997). □

Se o movimento de um sistema integrável é limitado, a região acessível do espaço de fase é finita e $\mathcal{M}_C$ é uma variedade compacta. Pelo item (c) do teorema, $\mathcal{M}_C$ é descrita por n ângulos, logo coincide topologicamente com um toro n-dimensional (ver Fig. 9.4). Usando as variáveis canônicas definidas acima, as equações $J_k = C_k$ definem em que toro o sistema se encontra, ao passo que os φs são coordenadas sobre o toro. Uma órbita que comece num desses n-toros nele permanece para sempre, por isso eles são chamados de *toros invariantes*.

De acordo com as equações (9.6.15), pequenos desvios das condições iniciais de um sistema integrável crescem linearmente com o tempo, e diz-se que o movimento de sistemas integráveis é *regular*. Já os sistemas não-integráveis podem, em certas regiões do espaço de fase, apresentar uma dependência ultra-sensível das condições iniciais: pequenos desvios das condições iniciais podem crescer exponencialmente com o tempo, tornando impossíveis previsões quanto ao comportamento do sistema mesmo a curto prazo. Tal tipo de movimento é dito *irregular* ou *caótico*. Assim, *caos* é uma propriedade privativa dos sistemas não-integráveis.

Na mecânica estatística, a fim de evitar a necessidade do cálculo de médias temporais que exigem o conhecimento da solução exata das equações de movimento, uma suposição fundamental é feita – a famosa *hipótese ergódica*. Segundo essa hipótese, no curso do tempo o sistema explora toda a região do espaço de fase energeticamente disponível (a chamada *superfície de energia*) e acaba por cobrir essa região uniformemente. Isto permite substituir médias temporais por médias sobre a superfície de energia no espaço de fase. No caso de um sistema integrável, a existência de n constantes de movimento

[5] No presente contexto, compacto significa a mesma coisa que fechado e limitado. Um conjunto é conexo se *não* é a união de conjuntos disjuntos.

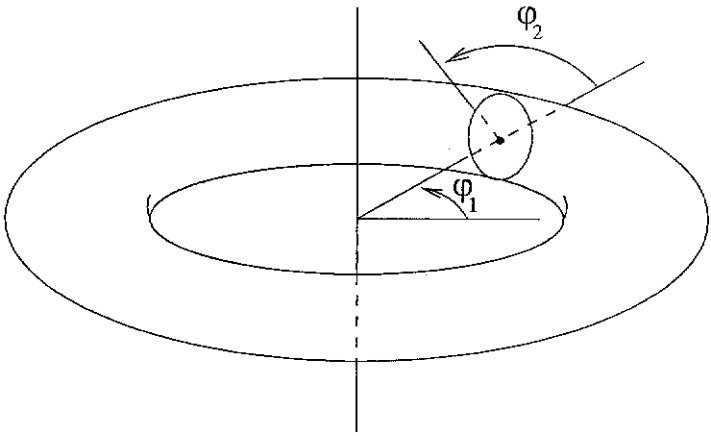


Figura 9.4: Toro invariante no caso n=2.

independentes implica que cada trajetória do sistema pode explorar no máximo uma variedade n-dimensional no espaço de fase de dimensão $2n$. Exceto no caso trivial em que $n = 1$, tal variedade é menor do que a superfície de energia, cuja dimensionalidade é $2n - 1$. Portanto, a hipótese ergódica é falsa em geral.

Uma pergunta natural é a seguinte: a integrabilidade constitui a regra ou a exceção? Os problemas que estão resolvidos exatamente nos textos de mecânica são todos integráveis, o que pode provocar a falsa impressão de que os sistemas integráveis são genéricos. Sabe-se que isto não é verdade. Os sistemas integráveis são os que costumam ser considerados nos livros de mecânica pelo simples fato de serem analiticamente solúveis, mas constituem menos a regra do que a exceção. Um exemplo importante de não-integrabilidade é caso geral do célebre problema dos três corpos da mecânica celeste. O surpreendente, no entanto, é que a não-integrabilidade pode se manifestar em sistemas mecânicos tão simples quanto o pêndulo duplo, que tem apenas dois graus de liberdade (Shinbrot *et al.* 1992; Hand & Finch 1998).

Uma questão relacionada, mas um pouco diferente, é se a integrabilidade resiste a pequenas perturbações. Assim como George Birkhoff e Enrico Fermi, muitos matemáticos e físicos teóricos acreditavam que a mais ligeira perturbação de um sistema integrável destruiria a integrabilidade, o que tenderia a validar a hipótese ergódica e, conseqüentemente, a mecânica estatística. O que ocorre, no entanto, é que em sua "maioria" os toros sobrevivem a uma perturbação, embora em forma distorcida. Se a perturbação é suficientemente pequena, os toros que são apenas deformados preenchem a maior parte do espaço de fase. Os toros destruídos distribuem-se entre os que são preservados de uma maneira extremamente irregular e o volume que eles ocupam no espaço de fase tende a zero juntamente com a perturbação. Isto foi demonstrado rigorosamente por Vladimir I. Arnold e Jürgen Moser, independentemente, em 1962-63, com base em sugestões feitas por A. N. Kolmogorov em 1954, e tornou-se conhecido como teorema de Kolmogorov-Arnold-Moser ou, abreviadamente, "teorema KAM".

A variedade em que se desenrola o movimento de um sistema integrável com hamiltoniana $H_0(J)$ é um n-toro T^n. Suponhamos que as freqüências $\omega_1, \ldots, \omega_n$ sejam incomensuráveis, ou, mais precisamente, que elas sejam racionalmente independentes, isto é, dados n números inteiros $m_1, \ldots, m_n$ tenhamos

$$\sum_{k=1}^{n} m_k \omega_k = 0 \iff m_1 = \cdots = m_n = 0 \ . \tag{9.6.17}$$

Depois de transcorrido um tempo suficientemente longo, a órbita retorna a uma vizinhança arbitrariamente pequena do ponto de partida sobre T^n mas não se fecha: dizemos que as órbitas são *densas* em T^n. O movimento é dito *quase-periódico* e o toro percorrido pela trajetória do sistema é chamado de *toro não-ressonante*. Quando as freqüências são racionalmente dependentes os toros são ditos *toros ressonantes*. Consideremos agora a hamiltoniana perturbada

$$H(\varphi, J; \epsilon) = H_0(J) + \epsilon H_1(\varphi, J) \tag{9.6.18}$$

e suponhamos que a perturbação quebre a integrabilidade do sistema. O problema fundamental consiste em determinar se, e em que sentido, este sistema é estável: a perturbação modifica apenas ligeiramente os movimentos do sistema descrito por $H_0(J)$ ou os destrói completamente? Argumentos genéricos e exemplos específicos (Arnold 1976; Ott 1993) indicam que, tipicamente, os toros ressonantes são destruídos por menor que seja a perturbação. No que se refere aos toros não-ressonantes, a pergunta é respondida em larga medida pelo teorema de Kolmogorov, Arnold e Moser.

Teorema KAM. Se as freqüências de um sistema hamiltoniano integrável H_0 são racionalmente independentes e suficientemente irracionais,[6] então, para ϵ suficientemente pequeno, as soluções do sistema perturbado (9.6.18) são preponderantemente quase-periódicas e só diferem ligeiramente das do sistema não-perturbado. Em sua maioria, os toros não-ressonantes de H_0 são apenas levemente deformados e o sistema perturbado também possui toros não-ressonantes, sobre os quais as órbitas são densas.

O teorema foi enunciado numa forma um tanto vaga e qualitativa para não afogar o leitor em tecnicalidades. Para a formulação precisa e a demonstração rigorosa, que, como assinala Berry, é "longa, intrincada e sutil", o leitor interessado é remetido aos trabalhos originais de Kolmogorov, Arnold e Moser, citados em José & Saletan (1998). Demonstrações simplificadas – mas ainda assim muito difíceis – de casos particulares do teorema encontram-se em Thirring (1997) e Chandre & Jauslin (1998). As idéias essenciais da demonstração estão expostas com preocupação didática em Berry (1978), Ozorio de Almeida (1995) e, sobretudo, em José & Saletan (1998).

O teorema KAM é uma grande proeza matemática, mas sua relevância física é considerada até certo ponto duvidosa. Infelizmente, a estabilidade do sistema solar não fica inequivocamente estabelecida porque o teorema parece requerer perturbações extremamente pequenas, muito menores do que as indicadas por experiências numéricas (Tabor 1989), além de ser omisso quanto ao destino dos movimentos com freqüências racionalmente dependentes. As órbitas às quais o teorema não se aplica estão densamente distribuídas no conjunto das órbitas abrangidas pelo teorema, de modo que os sistemas hamiltonianos em geral apresentam uma mistura complicada de movimentos regulares

[6] É verdade, alguns números são mais irracionais do que outros! Diz-se que as n freqüências racionalmente independentes $\omega_1, \ldots, \omega_n$ são suficientemente irracionais se existem números reais positivos $\alpha > n$ e β tais que é satisfeita a chamada *condição diofantina* $|\sum m_k \omega_k| \geq \beta |\mathbf{m}|^{-\alpha}$, onde $|\mathbf{m}| = |m_1| + \cdots + |m_n|$, qualquer que seja o vetor $\mathbf{m} = (m_1, \ldots, m_n) \neq 0$ cujas componentes são números inteiros.

com movimentos irregulares ou caóticos. As chamadas *lacunas de Kirkwood* (vazios no cinturão de asteróides entre Marte e Júpiter a distâncias do Sol em que os períodos de revolução não-perturbados do asteróide e de Júpiter são comensuráveis), descobertas em 1866, foram a princípio interpretadas com base no teorema KAM como resultantes da destruição de toros invariantes na vizinhança das freqüências ressonantes (Berry 1978). Esta explicação, no entanto, não é considerada conclusiva nem a mais plausível. Cálculos recentes parecem indicar que as lacunas de Kirkwood devem ser atribuídas ao comportamento caótico de sistemas hamiltonianos (Wisdom 1987; Lissauer 1999).

9.7 Variáveis de Ação no Problema de Kepler

Pelo papel que desempenhou no desenvolvimento da mecânica quântica, o tratamento do problema do movimento ligado de uma partícula sujeita a uma força central coulombiana pelo método das variáveis de ação-ângulo é de considerável interesse histórico. O potencial

$$V(r) = -\frac{k}{r} \quad , \quad k > 0 \ , \tag{9.7.1}$$

é da forma considerada no Exemplo 9.3.1, com $a(r) = -k/r$ e $b(\theta) = 0$. Portanto o sistema é separável e multiperiódico, de modo que as variáveis de ação escrevem-se

$$J_\varphi = \frac{1}{2\pi}\oint p_\varphi d\varphi = \frac{\alpha_\varphi}{2\pi}\int_0^{2\pi} d\varphi = \alpha_\varphi \ , \tag{9.7.2}$$

$$J_\theta = \frac{1}{2\pi}\oint p_\theta d\theta = \frac{1}{2\pi}\oint \frac{\partial W}{\partial \theta} d\theta = \frac{1}{2\pi}\oint \sqrt{\alpha_\theta^2 - \alpha_\varphi^2/\operatorname{sen}^2\theta}\, d\theta \ , \tag{9.7.3}$$

$$J_r = \frac{1}{2\pi}\oint p_r dr = \frac{1}{2\pi}\oint \frac{\partial W}{\partial r} dr = \frac{1}{2\pi}\oint \sqrt{2m\alpha_1 + \frac{2mk}{r} - \frac{\alpha_\theta^2}{r^2}}\, dr \ , \tag{9.7.4}$$

onde usamos as equações (9.3.18) e (9.3.19). As integrais (9.7.3) e (9.7.4) podem ser elegantemente calculadas usando técnicas de integração no plano complexo (Goldstein 1980; Saletan & Cromer 1971), mas vamos preferir fazê-lo por métodos elementares.

Necessariamente $\alpha_\theta^2 - \alpha_\varphi^2/\operatorname{sen}^2\theta \geq 0$, de modo que $\theta_0 \leq \theta \leq \pi - \theta_0$ onde $\operatorname{sen}\theta_0 = \alpha_\varphi/\alpha_\theta$. Para o cálculo da integral (9.7.3) devemos tomar a raiz quadrada positiva ($p_\theta > 0$) na região em que θ cresce de θ_0 a $\pi - \theta_0$ e raiz quadrada negativa ($p_\theta < 0$) na região em que θ decresce de $\pi - \theta_0$ a θ_0, fechando o circuito. Assim,

$$J_\theta = \frac{1}{2\pi} 2\int_{\theta_0}^{\pi-\theta_0} \sqrt{\alpha_\theta^2 - \alpha_\varphi^2/\operatorname{sen}^2\theta}\, d\theta = \frac{\alpha_\theta}{\pi}\int_{\theta_0}^{\pi-\theta_0} \left(1 - \frac{\operatorname{sen}^2\theta_0}{\operatorname{sen}^2\theta}\right)^{1/2} d\theta \ . \tag{9.7.5}$$

A fim de tornar simétrico o intervalo de integração, é conveniente introduzir a nova variável ψ definida por $\theta = \psi + \pi/2$ e o ângulo $\psi_0 = \pi/2 - \theta_0$, de tal maneira que

$$J_\theta = \frac{\alpha_\theta}{\pi}\int_{-\psi_0}^{\psi_0} \left(1 - \frac{\cos^2\psi_0}{\cos^2\psi}\right)^{1/2} d\psi = \frac{2\alpha_\theta}{\pi}\int_0^{\psi_0} \frac{\sqrt{\operatorname{sen}^2\psi_0 - \operatorname{sen}^2\psi}}{\cos\psi} d\psi \ , \tag{9.7.6}$$

onde

$$\cos\psi_0 = \operatorname{sen}\theta_0 = \alpha_\varphi/\alpha_\theta \ . \tag{9.7.7}$$

A forma do integrando em (9.7.6) sugere a mudança de variáveis

$$\operatorname{sen}\psi = \operatorname{sen}\psi_0 \operatorname{sen} u \ , \tag{9.7.8}$$

que resulta em

$$J_\theta = \frac{\alpha_\theta}{\pi}\int_0^{\pi/2} \frac{\operatorname{sen}^2\psi_0 \cos^2 u}{1-\operatorname{sen}^2\psi_0 \operatorname{sen}^2 u}du = \frac{\alpha_\theta}{\pi}\int_0^{\pi/2}\Big(1-\frac{\cos^2\psi_0}{1-\operatorname{sen}^2\psi_0 \operatorname{sen}^2 u}\Big)du \ . \tag{9.7.9}$$

Com o emprego da fórmula (Gradshteyn & Ryzhik 1980)

$$\int \frac{dx}{1-k^2\operatorname{sen}^2 x} = (1-k^2)^{-1/2}\tan^{-1}\big(\sqrt{1-k^2}\tan x\big) \quad , \quad |k|<1 \ ,$$

obtemos, finalmente,

$$J_\theta = \alpha_\theta(1-\cos\psi_0) = \alpha_\theta \ - \alpha_\varphi \ . \tag{9.7.10}$$

Quanto à variável de ação radial, temos

$$J_r = \frac{1}{2\pi}2\int_{r_-}^{r_+}\sqrt{2m\alpha_1 + \frac{2mk}{r} - \frac{\alpha_\theta^2}{r^2}}\,dr = \frac{1}{\pi}\int_{r_-}^{r_+}\frac{\sqrt{2m\alpha_1 r^2 + 2mkr - \alpha_\theta^2}}{r}\,dr \ , \tag{9.7.11}$$

onde r_+ e r_- são as raízes de

$$2m\alpha_1 + \frac{2mk}{r} - \frac{\alpha_\theta^2}{r^2} = 0 \ . \tag{9.7.12}$$

Evidentemente, estamos considerando $E<0$, pois apenas neste caso o movimento é periódico e a variável radial é limitada. Escrevendo $\alpha_1 = E = -\mathcal{E}$ com $\mathcal{E}>0$, as soluções de (9.7.12) são

$$r_\pm = \frac{k \pm \sqrt{k^2 - 2\mathcal{E}\alpha_\theta^2/m}}{2\mathcal{E}} \ . \tag{9.7.13}$$

Usando a fórmula (Gradshteyn & Ryzhik 1980)

$$\int \frac{\sqrt{a+bx+cx^2}}{x}\,dx = \sqrt{a+bx+cx^2} - \sqrt{-a}\,\operatorname{sen}^{-1}\Big(\frac{2a+bx}{x\sqrt{\Delta}}\Big) - \frac{b}{2\sqrt{-c}}\operatorname{sen}^{-1}\Big(\frac{2cx+b}{\sqrt{\Delta}}\Big) \ ,$$

válida para $a < 0$, $c < 0$ e $\Delta = b^2 - 4ac > 0$, com um pouco de trabalho algébrico deduz-se

$$J_r = -\alpha_\theta + \left(-\frac{2m}{E}\right)^{1/2} \frac{k}{2} \ . \tag{9.7.14}$$

Combinando este resultado com as Eqs.(9.7.2) e (9.7.10) obtém-se

$$H = E = -\frac{mk^2/2}{(J_r + J_\theta + J_\varphi)^2} \ , \tag{9.7.15}$$

donde

$$\omega_r = \frac{\partial H}{\partial J_r} = \frac{mk^2}{(J_r + J_\theta + J_\varphi)^3} = \left(-\frac{8E^3}{mk^2}\right)^{1/2} \ . \tag{9.7.16}$$

Há uma degenerescência completa, pois as demais freqüências ω_θ e ω_φ são iguais a ω_r. Sendo iguais entre si, as três freqüências são obviamente comensuráveis. Isto inevitavelmente teria que ocorrer em virtude de o movimento da partícula ser periódico, já que para $E < 0$ as órbitas são fechadas (circunferências ou elipses).

9.8 Invariantes Adiabáticos

Consideremos um sistema mecânico com um parâmetro característico λ e suponhamos que, sob a influência de agentes externos, λ varie muito lentamente ("adiabaticamente"). Por exemplo, uma massa m suspensa por um fio de comprimento l pode estar executando pequenas oscilações com um período τ. Se l for gradualmente reduzido, puxando-se o fio através de um buraco no teto, de modo que variações sensíveis de l só ocorram numa escala de tempo $T \gg \tau$, dizemos que o pêndulo está sofrendo uma transformação adiabática. Uma grandeza física que permanece constante durante uma transformação adiabática é dita um *invariante adiabático*.

Mostremos que, para um oscilador harmônico de freqüência lentamente variável, a grandeza E/ν é um invariante adiabático. A energia total de um oscilador harmônico é

$$E = \frac{m}{2}\dot{x}^2 + \frac{m\omega^2}{2}x^2 \ . \tag{9.8.1}$$

A freqüência angular ω poderia estar variando lentamente porque $\omega^2 = g/l$, como no caso do pêndulo, com l variando adiabaticamente. Qualquer que seja o mecanismo, podemos escrever

$$\frac{dE}{dt} = m\dot{x}\ddot{x} + m\omega^2 x\dot{x} + m\omega x^2 \frac{d\omega}{dt} \ . \tag{9.8.2}$$

Esta equação contém termos rápidos e termos lentos. Durante um intervalo de tempo necessário para que os termos lentos E e ω variem significativamente, os termos rápidos executam um grande número de oscilações, de modo que podemos substituí-los por seus valores médios num período de oscilação. Assim procedendo, ficamos com

$$\frac{dE}{dt} = \langle m\dot{x}\ddot{x} + m\omega^2 x\dot{x}\rangle + \langle m\omega x^2\rangle \frac{d\omega}{dt} \ . \tag{9.8.3}$$

A média do primeiro termo num ciclo é zero e $\langle m\omega x^2\rangle = E/\omega$, donde

$$\frac{dE}{E} = \frac{d\omega}{\omega} \quad \Longrightarrow \quad \frac{E}{\omega} = \text{constante} \ . \tag{9.8.4}$$

Levando em conta que $\nu = \omega/2\pi$ podemos escrever

$$\frac{E}{\nu} = \text{invariante adiabático} \ , \tag{9.8.5}$$

como pretendíamos demonstrar.

Inspecionando as Eqs.(9.5.21) e (9.5.24) verifica-se que E/ω é exatamente a variável de ação J do oscilador, de modo que, para um oscilador harmônico unidimensional, a variável de ação é um invariante adiabático. Desejamos mostrar, a seguir, que isto não é uma peculiaridade do oscilador, mas uma propriedade geral das variáveis de ação. Seja $H(q,p;\lambda)$ a hamiltoniana do sistema, suposto unidimensional. Quando λ é constante, sejam (J,φ) as variáveis de ação-ângulo tais que $H = H(J;\lambda)$. É conveniente tomar para a transformação canônica $(q,p) \rightarrow (\varphi, J)$ uma função geradora do primeiro tipo $W^*(q,\varphi;\lambda)$, que está relacionada a $W(q,J;\lambda)$ por

$$W^*(q,\varphi;\lambda) = W(q,J;\lambda) - \varphi J \ , \tag{9.8.6}$$

como se depreende de (8.1.13a). Quando λ varia com o tempo, W^* adquire uma dependência temporal explícita e a hamiltoniana transformada é

$$K(\varphi,J;\lambda) = H(J;\lambda) + \frac{\partial W^*}{\partial t} = H(J;\lambda) + \dot{\lambda}\frac{\partial W^*}{\partial \lambda} \ . \tag{9.8.7}$$

Agora a variável de ação J não é mais constante de movimento, mas satisfaz a equação de Hamilton

$$\dot{J} = -\frac{\partial K}{\partial \varphi} = -\dot{\lambda}\frac{\partial^2 W^*}{\partial\varphi\partial\lambda} \ . \tag{9.8.8}$$

No caso de uma perturbação adiabática, $\dot{\lambda}$ é uma quantidade muito pequena, de modo que J também varia lentamente. As variáveis lentas λ e J mudam pouco durante um período de oscilação do sistema não perturbado. Assim como fizemos no problema do oscilador, podemos obter uma versão aproximada para a equação (9.8.8) substituindo o coeficiente de $\dot{\lambda}$ em (9.8.8) por seu valor médio num ciclo do sistema não perturbado, resultando[7]

$$\dot{J} = -\dot{\lambda}\left\langle \frac{\partial^2 W^*}{\partial\varphi\partial\lambda}\right\rangle \ . \tag{9.8.9}$$

[7] A substituição de (9.8.8) por (9.8.9) costuma ser chamada de *princípio da média* (Arnold 1976).

Da generalização das equações (9.2.11) e (9.2.12) para um potencial arbitrário $V(q;\lambda)$ vemos que W exprime-se por meio da integral indefinida

$$W = \int p\, dq \ , \tag{9.8.10}$$

de modo que num ciclo do sistema não perturbado W aumenta de $2\pi J$. Ao mesmo tempo, φJ também aumenta de $2\pi J$ porque, de acordo com a Eq.(9.5.15), φ cresce de 2π em cada ciclo completo, ao passo que J permanece constante. Portanto, a Eq.(9.8.6) assegura que W^* é uma função periódica de φ com período 2π. O mesmo ocorre com sua derivada parcial em relação a λ, que, portanto, admite a expansão em série de Fourier

$$\frac{\partial W^*}{\partial \lambda} = \sum_{m=-\infty}^{\infty} A_m(J;\lambda) e^{im\varphi} \ . \tag{9.8.11}$$

Conseqüentemente,

$$\left\langle \frac{\partial^2 W^*}{\partial \varphi \partial \lambda} \right\rangle = \left\langle \sum_{m\neq 0} imA_m(J;\lambda) e^{im\varphi} \right\rangle = \sum_{m\neq 0} imA_m(J;\lambda) \langle e^{im\varphi} \rangle = 0 \ , \tag{9.8.12}$$

porque o termo constante está ausente da soma em (9.8.12) e todos os outros termos têm média zero num ciclo completo. Levando este resultado em (9.8.9) resulta

$$\dot{J} = 0 \ , \tag{9.8.13}$$

o que estabelece a invariância adiabática da variável de ação.[8]

■ **Exercício 9.8.1.** Uma partícula desliza sobre uma superfície horizontal sem atrito, oscilando livremente entre duas paredes verticais separadas por uma distância L. Se as colisões da partícula com as paredes são elásticas e a distância entre as paredes é variada lentamente, prove que $EL^2 =$ invariante adiabático. ■

Os invariantes adiabáticos encontram aplicação numa extensa gama de situações de interesse físico (Gleiser & Kozameh 1980; Crawford 1990). Especialmente importantes são os problemas envolvendo movimento de uma partícula num campo eletromagnético externo, cujos invariantes adiabáticos associados ajudam a explicar certos fenômenos observados em astrofísica e física de plasmas, além de desempenharem um papel significativo na teoria dos aceleradores de partículas (Crawford 1990).

9.9 Teoria de Hamilton-Jacobi e Mecânica Quântica

Na época da chamada "velha teoria quântica", fase de transição entre a descoberta do quantum de ação por Planck (1900) e a formulação da mecânica quântica moderna por Heisenberg, Born, Jordan, Dirac e Schrödinger (1925-26), o método mais geral então

[8] O princípio da média pode ser justificado pelo teorema da média, conferindo pleno rigor matemático aos argumentos que conduzem à invariância adiabática das variáveis de ação (Arnold 1976).

conhecido de obter os níveis de energia de sistemas microscópicos consistia em impor as chamadas condições quânticas

$$J_i = \frac{1}{2\pi} \oint p_i dq_i = n_i \hbar \ , \tag{9.9.1}$$

onde os n_i são números inteiros e $\hbar$ é a constante de Planck dividida por 2π. Propostas em 1915, as condições (9.9.1), conhecidas como regras de quantização de Wilson-Sommerfeld ou Bohr-Sommerfeld, davam bons resultados para sistemas simples como o oscilador harmônico, o rotor rígido e o átomo de hidrogênio, este último mesmo com a inclusão de correções relativísticas.[9] No caso do átomo de hidrogênio, por exemplo, escrevendo $k = e^2$ em unidades eletrostáticas CGS, impondo

$$J_r = n_r \hbar \ , \quad J_\theta = n_\theta \hbar \ , \quad J_\varphi = n_\varphi \hbar \ , \tag{9.9.2}$$

e usando (9.7.15) encontram-se os níveis de energia corretos

$$E_n = -\frac{me^4}{2\hbar^2 n^2} \ , \quad n = 1, 2, \ldots \ , \tag{9.9.3}$$

onde $n = n_r + n_\theta + n_\varphi$.

Dentre a infinidade contínua de movimentos possíveis do elétron no átomo de hidrogênio, as condições quânticas (9.9.2) selecionavam um conjunto discreto de órbitas permitidas. Segundo a *hipótese adiabática* de Ehrenfest, "se um sistema é afetado de um modo adiabático reversível, movimentos permitidos são transformados em movimentos permitidos." Com base nesse princípio heurístico, e mostrando que sistemas mecânicos complicados podem ser adiabaticamente transformados em sistemas mecânicos simples, Ehrenfest explorou a invariância adiabática das variáveis de ação para justificar as regras de quantização de Bohr-Sommerfeld (van der Waerden 1966). Essas regras não podiam representar a última palavra, no entanto, porque fracassavam rotundamente quando aplicadas a átomos com mais de um elétron ou a moléculas diatômicas.

Analogia Entre Mecânica Clássica e Óptica Geométrica

A teoria de Hamilton-Jacobi desempenhou um papel crucial na construção da "mecânica ondulatória" por Schrödinger, apoiado na analogia entre a mecânica e a óptica geométrica desenvolvida por Hamilton no período de 1828 a 1837. Por simplicidade, limitaremos a discussão ao problema de uma partícula no potencial $V(x, y, z)$ em coordenadas cartesianas, embora a análise possa ser estendida a sistemas muito mais gerais (Schrödinger 1982). A equação de Hamilton-Jacobi em coordenadas cartesianas

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2\right] + V(x, y, z) + \frac{\partial S}{\partial t} = 0 \tag{9.9.4}$$

[9] Eis aqui um belo exemplo de um dos aspectos mais fascinantes da física teórica. Introduzidas originalmente no domínio do muito grande (mecânica celeste), as variáveis de ação revelaram-se igualmente importantes no domínio do muito pequeno (física atômica).

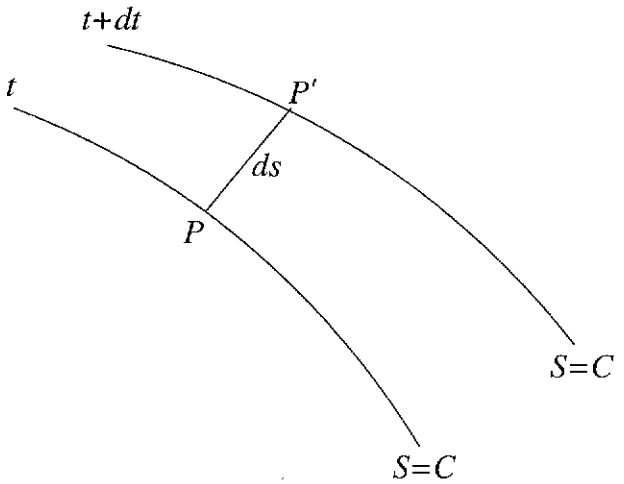


Figura 9.5: Frentes de onda ou superfícies de fase constante na analogia entre mecânica e óptica geométrica.

possui uma integral completa da forma

$$S(x,y,z,t) = W(x,y,z) - Et \ , \tag{9.9.5}$$

onde a função característica de Hamilton satisfaz

$$(\nabla W)^2 = 2m(E-V) \ , \tag{9.9.6}$$

sendo E a energia total (constante) da partícula. Consideremos, seguindo Hamilton, a família contínua de superfícies definidas pela equação

$$S(x,y,z,t) = W(x,y,z) - Et = C \ , \tag{9.9.7}$$

onde C é uma constante. De

$$\mathrm{p} = \nabla W \tag{9.9.8}$$

conclui-se que, em cada instante t, a trajetória da partícula é perpendicular à superfície definida por (9.9.7). Em outras palavras, pode-se obter todas as trajetórias mecânicas possíveis por meio da construção das trajetórias ortogonais às superfícies $S =$ constante. Isto sugere interpretar as trajetórias mecânicas do mesmo modo que os raios luminosos da óptica geométrica, sendo a ação S encarada como a fase de um processo ondulatório *fictício*. A velocidade de fase do referido processo pode ser determinada como se segue. Considere duas frentes de onda ou superfícies de fase constante nos instantes t e $t+dt$, respectivamente. Seja ds a distância ao longo da normal entre um ponto P da primeira superfície e o ponto correspondente P' da segunda superfície, como mostra a Fig. 9.5.

Como a ação S tem o mesmo valor nas duas frentes de onda, a equação (9.9.7) fornece $dW - Edt = 0$. Por outro lado, de acordo com (9.9.6),

$$Edt = dW = |\nabla W| ds = \sqrt{2m(E-V)}\, ds \ , \tag{9.9.9}$$

o que determina a velocidade de fase u como sendo

$$u = \frac{ds}{dt} = \frac{E}{\sqrt{2m(E-V)}} \ . \tag{9.9.10}$$

Mecânica Ondulatória e Equação de Schrödinger

Essa analogia entre mecânica e óptica geométrica induziu Erwin Schrödinger a pensar que se a mecânica clássica fosse apenas uma forma aproximada de uma mecânica com características ondulatórias, sua falha em escala microscópica poderia ser entendida como análoga ao fracasso da óptica geométrica na explicação dos fenômenos tipicamente ondulatórios de interferência e difração (Schrödinger 1982). Assim, em sua segunda comunicação da monumental série intitulada "Quantização Como um Problema de Autovalores", publicada em 1926, Schrödinger conjecturou que S deveria ser a fase de um processo ondulatório *verdadeiro* Ψ expresso por[10]

$$\Psi = \exp\Big(\frac{i}{\hbar}S\Big) = \exp\left\{\frac{i}{\hbar}\big[-Et + W(x,y,z)\big]\right\} \ . \tag{9.9.11}$$

Ora, um processo ondulatório com velocidade de fase u é tipicamente governado pela equação de onda

$$\nabla^2\Psi - \frac{1}{u^2}\frac{\partial^2\Psi}{\partial t^2} = 0 \ . \tag{9.9.12}$$

Substituindo nesta equação a velocidade de fase (9.9.10) e a função de onda Ψ da forma (9.9.11) resulta

$$-\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi = E\Psi \ , \tag{9.9.13}$$

que é a célebre equação de Schrödinger independente do tempo. Levando em conta (9.9.11), esta última equação é equivalente a

$$-\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi = i\hbar\frac{\partial\Psi}{\partial t} \ , \tag{9.9.14}$$

que é a equação de Schrödinger dependente do tempo.

Esses argumentos heurísticos, a despeito de seu grande poder de convencimento e irresistível beleza, não constituem uma dedução da equação de Schrödinger. Procedendo de forma inversa, isto é, *postulando* a Eq.(9.9.14) e escrevendo Ψ na forma $\Psi = \exp(iS/\hbar)$, verifica-se imediatamente que a equação obedecida por S é

[10] A presença da constante $\hbar = h/2\pi$ em (9.9.11) tem por finalidade garantir que a freqüência seja dada pela relação de Planck $E = h\nu$.

$$\frac{1}{2m}(\nabla S)^2 + V + \frac{\partial S}{\partial t} - \frac{i\hbar}{2m}\nabla^2 S = 0 \ . \tag{9.9.15}$$

No limite clássico $\hbar \to 0$ esta equação reduz-se à equação de Hamilton-Jacobi (9.9.4) e a fase da função de onda Ψ coincide com a ação clássica. Assim como o limite clássico da mecânica quântica na formulação de Heisenberg é a mecânica clássica expressa na linguagem dos parênteses de Poisson, o limite clássico da mecânica quântica na formulação de Schrödinger é a teoria de Hamilton-Jacobi. Incidentalmente, a Eq.(9.9.11) explica por que a equação de Hamilton-Jacobi é separável na forma de soma nos mesmos sistemas de coordenadas em que a equação de Schrödinger é separável na forma de produto.

Como $\mathbf{p} = \nabla S$, no limite clássico a trajetória da partícula é ortogonal à superfície de fase constante da função de onda Ψ. No limite clássico, portanto, os "raios" associados a Ψ são as trajetórias possíveis da partícula. Neste sentido, a mecânica clássica é um caso limite da mecânica quântica da mesma forma que a óptica geométrica é um caso limite da óptica física. É tentador imaginar que, se tivesse sido suficientemente audacioso, Hamilton poderia ter proposto a mecânica clássica como um autêntico caso limite de uma mecânica com características ondulatórias, antecipando em quase um século o advento da mecânica quântica. No entanto, a ausência de suporte experimental provavelmente teria levado os físicos da época a pôr de lado tal proposta como um mero delírio especulativo sem fundamento.

PROBLEMAS

9.1. Uma partícula move-se em uma dimensão no potencial $V(x) = k/x^2$, $k > 0$. Determine $x(t)$ pelo método de Hamilton-Jacobi se $x(0) = x_0$ e $\dot{x}(0) = 0$.

9.2. A lagrangiana de uma partícula em coordenadas cilíndricas (ρ, z, φ) é

$$L = \frac{m}{2}(\dot{\rho}^2 + \rho^2\dot{\varphi}^2 + \dot{z}^2) - V(\rho, \varphi, z) \ .$$

Introduza as *coordenadas parabólicas* (ξ, η, φ) definidas por $\rho = \sqrt{\xi\eta}$, $z = (\xi - \eta)/2$ com $\xi \geq 0$ e $\eta \geq 0$. Mostre que, inversamente, $\xi = r + z$, $\eta = r - z$, onde $r = \sqrt{\rho^2 + z^2}$. Prove que a hamiltoniana em coordenadas parabólicas tem a forma

$$H = \frac{2}{m}\frac{\xi p_\xi^2 + \eta p_\eta^2}{\xi + \eta} + \frac{p_\varphi^2}{2m\xi\eta} + V(\xi, \eta, \varphi) \ .$$

Se

$$V = \frac{a(\xi) + b(\eta)}{\xi + \eta} \ ,$$

com $a(\xi)$ e $b(\eta)$ funções arbitrárias, mostre que a equação de Hamilton-Jacobi independente do tempo é separável e obtenha uma integral completa. Considere o efeito Stark, isto é, a ação de um campo elétrico uniforme sobre um átomo de hidrogênio. Desprezando o movimento do próton e escolhendo o eixo z paralelo a $\mathbf{E}$, mostre que $V = -k/\sqrt{\rho^2 + z^2} - eEz$, $k > 0$. Comprove que, expresso em coordenadas parabólicas, este potencial é da forma acima e identifique as funções $a(\xi)$ e $b(\eta)$. Finalmente, utilizando a integral completa obtida para a equação de Hamilton-Jacobi, reduza a quadraturas a solução das equações de movimento do elétron.

9.3. Considere o problema de Kepler no plano, isto é, uma partícula de massa m no potencial $V(r) = k/r$, onde $r = \sqrt{x^2 + y^2}$. Obtenha a equação de Hamilton-Jacobi em termos das variáveis $u = r + x$, $v = r - x$ e encontre uma integral completa. Determine a equação da órbita da partícula em termos das coordenadas (u, v).

9.4. Uma partícula com energia total positiva move-se ao longo de uma linha reta sob a influência do potencial $V(x) = F|x|$, onde F é uma constante positiva. Use variáveis de ação-ângulo para encontrar o período do movimento como função da energia. Qual é o espectro energético resultante da aplicação das regras de Wilson-Sommerfeld?

9.5. Seja $H(q, p, t)$ a hamiltoniana de um sistema com n graus de liberdade e suponha que S satisfaça a equação diferencial parcial

$$H\Big(-\frac{\partial S}{\partial p_1}, \ldots, -\frac{\partial S}{\partial p_n}, p_1, \ldots, p_n, t\Big) + \frac{\partial S}{\partial t} = 0 \ ,$$

conhecida como equação de Hamilton-Jacobi na representação p. Que tipo de função geradora S deve ser considerada para se chegar a esta equação? Escreva as equações e indique o procedimento que permite obter a solução das equações de movimento uma vez conhecida uma integral completa $S(p, \alpha, t)$ da equação acima. Aplique esta formulação ao problema unidimensional de uma partícula sujeita a uma força constante, cuja

hamiltoniana é $H = p^2/2m - Fq$, e determine o movimento $q(t)$ da partícula. Por que a equação de Hamilton-Jacobi na representação p é, em geral, mais difícil de resolver do que na representação usual?

9.6. Uma partícula de carga e move-se no plano (x, y) na presença de um um campo magnético constante perpendicular ao plano. (i) Mostre que uma hamiltoniana que descreve este sistema é

$$H = \frac{1}{2m}p_x^2 + \frac{1}{2m}\Big(p_y - \frac{eB}{c}x\Big)^2$$

e obtenha a solução $x(t), y(t), p_x(t), p_y(t)$ das equações de movimento pelo método de Hamilton-Jacobi. (ii) Mostre que uma hamiltoniana alternativa para este sistema é

$$H = \frac{1}{2m}\Big(p_x + \frac{eB}{2c}y\Big)^2 + \frac{1}{2m}\Big(p_y - \frac{eB}{2c}x\Big)^2 \ .$$

Constate que a equação de Hamilton-Jacobi associada não é separável na forma usual. Isto mostra que a separabilidade da equação de Hamilton-Jacobi depende não apenas da escolha do sistema de coordenadas, mas também do calibre em que ela é escrita. Separe as variáveis fazendo $W = Cxy + \alpha_y y + X(x)$, onde α_y é uma constante arbitrária e C é uma constante convenientemente escolhida, e resolva as equações de movimento. Compare com os resultados obtidos no item (i).

9.7. Suponha que num problema com um grau de liberdade o potencial dependa linearmente do tempo, de modo que

$$H = \frac{p^2}{2m} - Atx \ ,$$

onde A é uma constante. Resolva as equações de movimento deste problema pelo método de Hamilton-Jacobi.

9.8. Mostre que a equação de Hamilton-Jacobi (9.2.1) para uma partícula livre unidimensional é separável na forma de produto, isto é, admite soluções da forma $S(q,t) = W(q)T(t)$. Usando esta técnica de separação de variáveis, obtenha uma integral completa e, a partir dela, a solução geral da equação de movimento para $q(t)$.

9.9. Uma partícula de massa m move-se ao longo de uma linha reta sujeita ao potencial $V(x) = V_0/\cos^2(x/\ell)$, $V_0 > 0$, $\ell > 0$. (i) Esboce o gráfico de V versus x para $x \in (-\ell\pi/2, \ell\pi/2)$. Mostre que para qualquer valor classicamente permitido da energia total o movimento da partícula é periódico. (ii) Use o método das variáveis de ação-ângulo para obter o período de oscilação em função da energia. (iii) Quais são os níveis de energia da partícula segundo as regras de quantização de Wilson-Sommerfeld?

9.10. Uma partícula num plano tem seu movimento descrito pela hamiltoniana

$$H = p_x p_y \cos\omega t + \frac{1}{2}(p_x^2 - p_y^2)\,\mathrm{sen}\,\omega t \ .$$

(i) Encontre uma integral completa da equação de Hamilton-Jacobi associada. (ii) Obtenha a solução geral das equações de movimento para $x(t)$ e $y(t)$.

9.11. O movimento de uma partícula num plano é determinado pela hamiltoniana

$$H = \frac{1}{2}\frac{p_x^2 + p_y^2}{x^2 + y^2} + \frac{k}{x^2 + y^2} \ ,$$

onde (x, y) são coordenadas cartesianas e $k > 0$. (i) Use a equação de Hamilton-Jacobi para obter a equação da trajetória da partícula. (ii) Prove que a órbita é uma seção cônica (elipse, parábola ou hipérbole).

9.12. Reconsidere a máquina de Atwood oscilante do Problema 1.5 no caso especial $M = 3m$. Mostre que a mudança de variáveis

$$\xi = \sqrt{r\left(1 + \operatorname{sen}\frac{\theta}{2}\right)} \quad , \quad \eta = \sqrt{r\left(1 - \operatorname{sen}\frac{\theta}{2}\right)}$$

com inversa

$$r = \frac{\xi^2 + \eta^2}{2} \quad , \quad \theta = 2\tan^{-1}\frac{\xi^2 - \eta^2}{2\xi\eta}$$

transforma a Lagrangiana dada no Problema 1.5 em

$$L = 2m(\xi^2 + \eta^2)(\dot{\xi}^2 + \dot{\eta}^2) - 2mg\frac{\xi^4 + \eta^4}{\xi^2 + \eta^2} \ ,$$

que dá lugar à hamiltoniana

$$H = \frac{1}{8m}\frac{p_\xi^2 + p_\eta^2}{\xi^2 + \eta^2} + 2mg\frac{\xi^4 + \eta^4}{\xi^2 + \eta^2} \ .$$

Mostre que a equação de Hamilton-Jacobi correspondente é separável e encontre uma integral completa (Tufillaro 1986).

9.13. Dê uma demonstração elementar do resultado do Exercício 9.8.1 calculando o trabalho realizado pela força externa para variar de dL a distância entre as paredes e levando em conta que a variação da energia da partícula deve ser igual ao trabalho realizado pela força aplicada. Sugestão: a força externa deve contrabalançar a força média que a partícula exerce sobre as paredes em virtude das colisões sucessivas.

9.14. Uma partícula de massa m move-se num plano sujeita ao campo gravitacional de duas massas muito maiores m_1 e m_2 que permanecem fixas sobre o eixo x nas posições $(\pm a, 0)$. Mostre que em *coordenadas elípticas* (u, v) definidas por

$$x = a\cosh u\cos v \quad , \quad y = a\operatorname{senh} u\operatorname{sen} v \ ,$$

a hamiltoniana deste problema é

$$H = \frac{1}{2ma^2}\frac{p_u^2 + p_v^2}{\cosh^2 u - \cos^2 v} + \frac{k}{\cosh u - \cos v} + \frac{k'}{\cosh u + \cos v} \ .$$

Mostre que a equação de Hamilton-Jacobi correspondente é separável e reduza a solução das equações de movimento a quadraturas.

9.15. Uma partícula com energia total negativa move-se em uma dimensão no potencial $V = -V_0/\cosh^2(\alpha x)$, onde α é uma constante positiva. Se V_0 varia lentamente com o tempo, prove que $(V_0 + E)/\sqrt{V_0}$ é um invariante adiabático.

9.16. Considere o sistema descrito no Problema 2.12 e suponha que o fio seja puxado de modo a reduzir lentamente o raio r da circunferência descrita com velocidade v pela partícula de massa m_2. Prove que o produto vr é um invariante adiabático. Apresente também uma demonstração elementar desse resultado, sem recorrer a variáveis de ação.

9.17. No desenvolvimento da analogia entre a mecânica clássica e a óptica geométrica, o leitor mais atento deve ter notado que a velocidade de fase (9.9.10) não coincide com a velocidade da partícula — é, na verdade, inversamente proporcional a ela. Esse descompasso parece inviabilizar a pretensão de encarar S como a fase de um processo ondulatório associado à partícula. No entanto, pode-se mostrar (Jackson 1999) que a velocidade de um pacote de ondas formado pela superposição de ondas numa faixa estreita de freqüências é a velocidade de grupo dada por $v_g = d\omega/dk$, onde k é o vetor número de onda. Partindo da expressão para $\omega(k)$ determinada por meio de (9.9.6), (9.9.8), da relação de Planck $E = h\nu$ e da relação de de Broglie $\mathbf{p} = \hbar\mathbf{k}$, prove que a velocidade de grupo é idêntica à velocidade da partícula.

9.18. Na mecânica quântica, a noção de *conjunto completo de observáveis comutativos* desempenha um papel importante. Seu análogo clássico, para um sistema com n graus de liberdade, é um conjunto $R_1(q,p), \ldots, R_n(q,p)$ de n funções independentes em involução. Prove o seguinte teorema (Sharan 1982): se $F(q,p)$ tem parêntese de Poisson nulo com todos os $R_k(q,p)$ então F é uma função dos R_k, isto é, existe uma função $\Phi : \boldsymbol{R}^n \to \boldsymbol{R}$ tal que $F(q,p) = \Phi(R_1(q,p), \ldots, R_n(q,p))$. Sugestão: conforme a demonstração do Lema 9.6.1, existe uma transformação canônica $(q,p) \to (Q,P)$ com $P_k = R_k(q,p)$.

Capítulo 10

TEORIA CLÁSSICA DE CAMPOS

> That one body may act upon another at a distance through a vacuum, without the mediation of any thing else, by and through which their action and force may be conveyed from one to the other, is to me so great an absurdity, that I believe no man who has in philosophical matters a competent faculty of thinking, can ever fall into it.
>
> Isaac Newton

Sistemas contínuos possuem um número infinito de graus de liberdade e são descritos por campos. É fato notável que praticamente todas as teorias de campos de interesse físico podem ser descritas pelos formalismos de Lagrange e Hamilton. As interações das partículas elementares, constituintes básicos da matéria, são expressas por meio de teorias quânticas de campos. Por sua vez, a construção das teorias quânticas das interações fundamentais da natureza depende crucialmente da possibilidade de primeiro formulá-las como teorias clássicas de campos nas linguagens lagrangiana e hamiltoniana, o que vem a ser o objeto do presente capítulo.

10.1 Teoria de Campos na Forma Lagrangiana

Um sistema mecânico com um número finito de graus de liberdade é descrito pelas coordenadas generalizadas $q_k(t)$. O sistema contínuo mais simples é descrito por uma coordenada $\varphi_{\mathrm{x}}(t)$ associada a cada ponto x do espaço, ou seja, o índice discreto k é substituído pelo índice contínuo x. Por simplicidade, consideraremos inicialmente campos em apenas uma dimensão espacial e, em vez de utilizar a coordenada espacial como subscrito, usaremos a notação tradicional $\varphi(x,t)$. Por exemplo, $\varphi(x,t)$ poderia representar o deslocamento transversal no instante t do ponto x de uma corda vibrante.

A lagrangiana de um sistema discreto envolve uma soma sobre todos os graus de liberdade, de modo que a lagrangiana de um sistema contínuo deve ser expressa em termos da integral espacial de uma função $\mathcal{L}$, chamada de *densidade lagrangiana*. A densidade lagrangiana deve conter um termo cinético, logo deve depender de $\dot{\varphi}(x,t) \equiv$

$\partial\varphi/\partial t$. Em contraste com a idéia de ação à distância, suporemos que o campo φ num ponto x interage consigo mesmo somente numa vizinhança infinitesimal desse ponto, de modo que $\mathcal{L}$ deve depender de $\varphi(x,t)$ e $\varphi(x+dx,t)$. Alternativamente, em vez desta última quantidade é melhor usar $\varphi'(x,t) \equiv \partial\varphi/\partial x$. Admitindo uma possível dependência explícita em x e t, a ação mais geral para uma teoria de campos unidimensional tem a forma

$$S = \int_{t_1}^{t_2} L dt = \int_{t_1}^{t_2} dt \int_{x_1}^{x_2} dx\, \mathcal{L}\Big(\varphi, \frac{\partial\varphi}{\partial x}, \frac{\partial\varphi}{\partial t}, x, t\Big) \ . \tag{10.1.1}$$

A equação de Lagrange para φ decorre do princípio de Hamilton

$$\delta S = \delta \int_{t_1}^{t_2} dt \int_{x_1}^{x_2} dx\, \mathcal{L} = 0 \ . \tag{10.1.2}$$

Como o exemplo da corda vibrante com extremos fixos sugere, a variação do campo deve anular-se não apenas nos extremos temporais mas também nos extremos espaciais:

$$\delta\varphi(x,t_1) = \delta\varphi(x,t_2) = 0 \quad , \quad \delta\varphi(x_1,t) = \delta\varphi(x_2,t) = 0 \ . \tag{10.1.3}$$

Executando a variação da ação (10.1.1) resulta

$$\delta S = \int_{t_1}^{t_2} dt \int_{x_1}^{x_2} dx \Big\{ \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\delta\dot{\varphi} + \frac{\partial\mathcal{L}}{\partial\varphi'}\delta\varphi' \Big\} \ . \tag{10.1.4}$$

Usando

$$\delta\dot{\varphi} = \frac{\partial(\delta\varphi)}{\partial t} \quad , \quad \delta\varphi' = \frac{\partial(\delta\varphi)}{\partial x} \ , \tag{10.1.5}$$

realizando integrações por partes e levando em conta (10.1.3), obtemos

$$\int_{t_1}^{t_2} dt \int_{x_1}^{x_2} dx \frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\delta\dot{\varphi} = \int_{x_1}^{x_2} dx \int_{t_1}^{t_2} dt \frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{\partial(\delta\varphi)}{\partial t} = \int_{x_1}^{x_2} dx \frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\delta\varphi\Big|_{t_1}^{t_2} - \int_{x_1}^{x_2} dx \int_{t_1}^{t_2} dt \frac{\partial}{\partial t}\Big(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\Big)\delta\varphi$$

$$= -\int_{t_1}^{t_2} dt \int_{x_1}^{x_2} dx \frac{\partial}{\partial t}\Big(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\Big)\delta\varphi \tag{10.1.6a}$$

e, analogamente,

$$\int_{t_1}^{t_2} dt \int_{x_1}^{x_2} dx \frac{\partial\mathcal{L}}{\partial\varphi'}\delta\varphi' = \int_{t_1}^{t_2} dt \frac{\partial\mathcal{L}}{\partial\varphi'}\delta\varphi\Big|_{x_1}^{x_2} - \int_{t_1}^{t_2} dt \int_{x_1}^{x_2} dx \frac{\partial}{\partial x}\Big(\frac{\partial\mathcal{L}}{\partial\varphi'}\Big)\delta\varphi$$

$$= -\int_{t_1}^{t_2} dt \int_{x_1}^{x_2} dx \frac{\partial}{\partial x}\Big(\frac{\partial\mathcal{L}}{\partial\varphi'}\Big)\delta\varphi \ . \tag{10.1.6b}$$

Substituindo estes resultados em (10.1.4), o princípio de Hamilton torna-se

$$\int_{t_1}^{t_2} dt \int_{x_1}^{x_2} dx \left\{ \frac{\partial\mathcal{L}}{\partial\varphi} - \frac{\partial}{\partial t}\Big(\frac{\partial\mathcal{L}}{\partial(\partial\varphi/\partial t)}\Big) - \frac{\partial}{\partial x}\Big(\frac{\partial\mathcal{L}}{\partial(\partial\varphi/\partial x)}\Big) \right\}\delta\varphi = 0 \ . \tag{10.1.7}$$

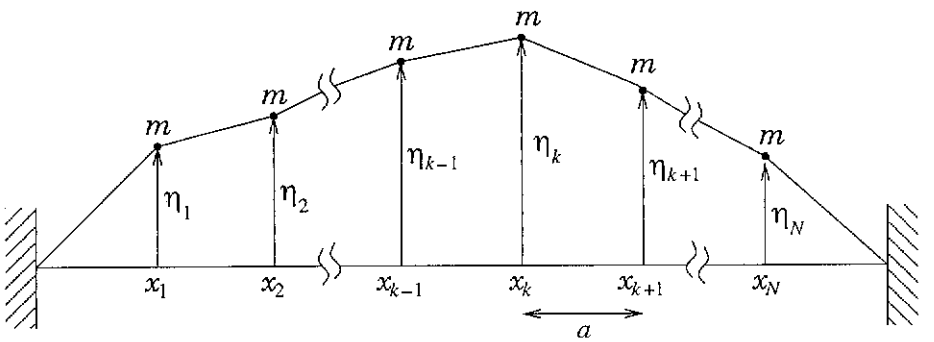


Figura 10.1: Modelo discreto para vibrações transversais de uma corda.

Pelo mesmo argumento utilizado na Seção 2.1, a arbitrariedade de $\delta\varphi$ implica a validade da equação de Lagrange

$$\frac{\partial}{\partial t}\Big(\frac{\partial \mathcal{L}}{\partial(\partial\varphi/\partial t)}\Big) + \frac{\partial}{\partial x}\Big(\frac{\partial \mathcal{L}}{\partial(\partial\varphi/\partial x)}\Big) - \frac{\partial \mathcal{L}}{\partial \varphi} = 0 \ . \tag{10.1.8}$$

■ **Exemplo 10.1.1.** Uma corda vibrante tem os extremos fixos $x = 0$ e $x = l$. Se σ é a densidade linear de massa e τ é a tensão na corda, mostrar que a densidade lagrangiana

$$\mathcal{L} = \frac{1}{2}\sigma\Big(\frac{\partial\varphi}{\partial t}\Big)^2 - \frac{1}{2}\tau\Big(\frac{\partial\varphi}{\partial x}\Big)^2 \tag{10.1.9}$$

gera a equação de movimento correta para pequenas vibrações transversais da corda.

Solução. Escrevendo $\mathcal{L} = \sigma\dot{\varphi}^2/2 - \tau\varphi'^2/2$ obtém-se imediatamente

$$\frac{\partial \mathcal{L}}{\partial(\partial\varphi/\partial t)} \equiv \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = \sigma\frac{\partial\varphi}{\partial t} \quad , \quad \frac{\partial \mathcal{L}}{\partial(\partial\varphi/\partial x)} \equiv \frac{\partial \mathcal{L}}{\partial \varphi'} = -\tau\frac{\partial\varphi}{\partial x} \quad , \quad \frac{\partial \mathcal{L}}{\partial \varphi} = 0 \ . \tag{10.1.10}$$

Levando estes resultados na equação de Lagrange (10.1.8) resulta

$$\sigma\frac{\partial^2\varphi}{\partial t^2} - \tau\frac{\partial^2\varphi}{\partial x^2} = 0 \ , \tag{10.1.11}$$

que é a conhecida equação de onda unidimensional da corda vibrante. ■

Transição do Discreto ao Contínuo

Certas teorias de campos (mas não todas) podem ser interpretadas como o limite contínuo de um sistema mecânico discreto com um número finito ou infinito enumerável

de graus de liberdade. A corda vibrante é um desses casos, e é instrutivo e esclarecedor encará-la como o limite contínuo de um sistema discreto conforme a descrição que se segue. Suponhamos que N massas iguais estejam conectadas por fios de massa desprezível, como representado na Fig. 10.1. Na situação de equilíbrio a tensão τ nos fios é constante e as massas estão igualmente espaçadas de a. Espera-se obter a corda contínua passando ao limite $a \to 0$ e $m \to 0$ de tal modo que a densidade linear de massa permaneça finita.

As coordenadas generalizadas são os deslocamentos transversais $\eta_1, \ldots, \eta_N$ com a restrição

$$\eta_0 = \eta_{N+1} = 0 \ , \tag{10.1.12}$$

que incorpora a condição de contorno de extremos fixos. A Fig. 10.2 e a segunda lei de Newton permitem escrever a equação de movimento da i-ésima partícula na forma

$$m\ddot{\eta}_i = \tau \operatorname{sen}\beta - \tau \operatorname{sen}\alpha \approx \tau(\tan\beta - \tan\alpha) \ , \tag{10.1.13}$$

supondo pequenos deslocamentos.[1] Usando

$$\tan\alpha = \frac{\eta_i - \eta_{i-1}}{a} \quad , \quad \tan\beta = \frac{\eta_{i+1} - \eta_i}{a} \ , \tag{10.1.14}$$

a equação (10.1.13) torna-se

$$m\ddot{\eta}_i = \frac{\tau}{a}\left[(\eta_{i+1} - \eta_i) - (\eta_i - \eta_{i-1})\right] \ . \tag{10.1.15}$$

Estas equações de movimento são geradas pela lagrangiana

$$L = \sum_{k=1}^{N} \frac{m}{2}\dot{\eta}_k^2 - \sum_{k=0}^{N} \frac{\tau}{2a}(\eta_{k+1} - \eta_k)^2 \ . \tag{10.1.16}$$

■ **Exercício 10.1.1.** Mostre que as equações de Lagrange geradas pela lagrangiana (10.1.16) são exatamente as Eqs.(10.1.15). ■

Pondo $\Delta x = a$, denotando por x a abscissa da k-ésima partícula e fazendo a mudança de notação

$$\eta_k(t) \to \varphi(x,t) \quad , \quad \eta_{k+1}(t) - \eta_k(t) \to \varphi(x + \Delta x, t) - \varphi(x,t) \ , \tag{10.1.17}$$

podemos reescrever a lagrangiana (10.1.16) na forma conveniente

[1] Justifica-se usar a tensão de equilíbrio τ em (10.1.13) porque ela já aparece multiplicada por termos de primeira ordem nas quantidades pequenas.

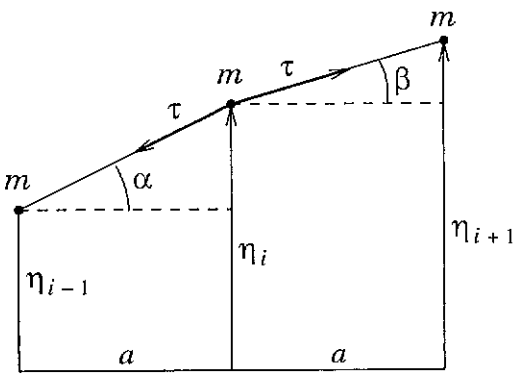


Figura 10.2: Diagrama de forças sobre a i-ésima partícula do modelo discreto para a corda vibrante.

$$L = \sum \Delta x \left[\frac{1}{2} \frac{m}{\Delta x} \left(\frac{\partial \varphi(x,t)}{\partial t} \right)^2 \right] - \sum \Delta x \frac{\tau}{2} \left(\frac{\varphi(x+\Delta x,t) - \varphi(x,t)}{\Delta x} \right)^2 . \qquad (10.1.18)$$

A passagem ao limite contínuo é efetuada fazendo $\Delta x \to 0$ e $m \to 0$ de tal modo que $m/\Delta x = \sigma$, resultando na lagrangiana

$$L = \int_0^l dx \left\{ \frac{\sigma}{2} \left(\frac{\partial \varphi}{\partial t} \right)^2 - \frac{\tau}{2} \left(\frac{\partial \varphi}{\partial x} \right)^2 \right\} , \qquad (10.1.19)$$

da qual se identifica imediatamente a densidade lagrangiana (10.1.9).

Vários Campos em Três Dimensões Espaciais

No caso de um sistema de N campos em três dimensões espaciais, representados coletivamente por $\varphi = (\varphi_1, \ldots, \varphi_N)$, as equações de Lagrange resultantes do princípio de Hamilton[2]

$$\delta S = \delta \int_\Omega d^4x \, \mathcal{L}\left(\varphi, \frac{\partial \varphi}{\partial t}, \nabla \varphi, \mathbf{x}, t \right) = 0 \qquad (10.1.20)$$

escrevem-se

$$\frac{\partial}{\partial t} \left(\frac{\partial \mathcal{L}}{\partial (\partial \varphi_\alpha / \partial t)} \right) + \nabla \cdot \left(\frac{\partial \mathcal{L}}{\partial (\nabla \varphi_\alpha)} \right) - \frac{\partial \mathcal{L}}{\partial \varphi_\alpha} = 0 \quad , \quad \alpha = 1, \ldots, N . \qquad (10.1.21)$$

A fim de demonstrar este resultado, consideremos

[2] Usamos a notação $d^3x \equiv dxdydz$ e $d^4x \equiv dtdxdydz$.

$$\delta S = \int_{t_1}^{t_2} dt \int_V d^3x \sum_{\alpha=1}^{N} \left\{ \frac{\partial \mathcal{L}}{\partial \varphi_\alpha} \delta\varphi_\alpha + \frac{\partial \mathcal{L}}{\partial \dot{\varphi}_\alpha} \delta\dot{\varphi}_\alpha + \frac{\partial \mathcal{L}}{\partial(\boldsymbol{\nabla}\varphi_\alpha)} \cdot \delta(\boldsymbol{\nabla}\varphi_\alpha) \right\} , \qquad (10.1.22)$$

onde as variações dos φ_α são mutuamente independentes e anulam-se nos extremos de integração temporal e na superfície que limita a região tridimensional V. Usando $\delta\dot{\varphi}_\alpha = \partial(\delta\varphi_\alpha)/\partial t$, uma integração por partes como no caso unidimensional fornece

$$\int_{t_1}^{t_2} dt \int_V d^3x \frac{\partial \mathcal{L}}{\partial \dot{\varphi}_\alpha} \delta\dot{\varphi}_\alpha = - \int_{t_1}^{t_2} dt \int_V d^3x \frac{\partial}{\partial t}\Big(\frac{\partial \mathcal{L}}{\partial \dot{\varphi}_\alpha}\Big)\delta\varphi_\alpha \ . \qquad (10.1.23)$$

Usando agora $\delta(\boldsymbol{\nabla}\varphi_\alpha) = \boldsymbol{\nabla}(\delta\varphi_\alpha)$ e lançando mão da identidade

$$\mathbf{A} \cdot \boldsymbol{\nabla} f = \boldsymbol{\nabla} \cdot (f\mathbf{A}) - f \boldsymbol{\nabla} \cdot \mathbf{A} \ , \qquad (10.1.24)$$

podemos efetuar uma integração por partes com a ajuda do teorema da divergência para obter

$$\begin{aligned} \int_{t_1}^{t_2} dt \int_V d^3x \frac{\partial \mathcal{L}}{\partial(\boldsymbol{\nabla}\varphi_\alpha)} \cdot \delta(\boldsymbol{\nabla}\varphi_\alpha) &= \int_{t_1}^{t_2} dt \int_V d^3x \frac{\partial \mathcal{L}}{\partial(\boldsymbol{\nabla}\varphi_\alpha)} \cdot \boldsymbol{\nabla}(\delta\varphi_\alpha) \\ &= \int_{t_1}^{t_2} dt \oint_\Sigma \mathbf{da} \cdot \frac{\partial \mathcal{L}}{\partial(\boldsymbol{\nabla}\varphi_\alpha)} \delta\varphi_\alpha - \int_{t_1}^{t_2} dt \int_V d^3x \, \boldsymbol{\nabla} \cdot \Big(\frac{\partial \mathcal{L}}{\partial(\boldsymbol{\nabla}\varphi_\alpha)}\Big)\delta\varphi_\alpha \\ &= - \int_{t_1}^{t_2} dt \int_V d^3x \, \boldsymbol{\nabla} \cdot \Big(\frac{\partial \mathcal{L}}{\partial(\boldsymbol{\nabla}\varphi_\alpha)}\Big)\delta\varphi_\alpha \ , \end{aligned} \qquad (10.1.25)$$

pois as variações $\delta\varphi_\alpha$ anulam-se na superfície Σ que limita a região espacial V. Introduzindo (10.1.23) e (10.1.25) em (10.1.22), o princípio de Hamilton toma a forma

$$\delta S = \int_{t_1}^{t_2} dt \int_V d^3x \sum_{\alpha=1}^{N} \Big\{ \frac{\partial \mathcal{L}}{\partial \varphi_\alpha} - \frac{\partial}{\partial t}\Big(\frac{\partial \mathcal{L}}{\partial \dot{\varphi}_\alpha}\Big) - \boldsymbol{\nabla} \cdot \Big(\frac{\partial \mathcal{L}}{\partial(\boldsymbol{\nabla}\varphi_\alpha)}\Big) \Big\} \delta\varphi_\alpha = 0 \ , \qquad (10.1.26)$$

donde resultam imediatamente as equações de Lagrange (10.1.21) porque as variações $\delta\varphi_\alpha$ são independentes e arbitrárias.

Campos Complexos

Certas teorias demandam o emprego de campos complexos. Um campo complexo equivale a dois campos reais independentes. Equivalentemente, em lugar das partes real e imaginária de um campo complexo ψ, podemos considerar o próprio campo e seu complexo conjugado ψ^* como campos independentes. Neste caso, para cada campo complexo teremos um par de equações de Lagrange da forma (10.1.21) tomando primeiro $\varphi_\alpha = \psi$ e em seguida $\varphi_\alpha = \psi^*$.

■ **Exercício 10.1.2.** Seja ψ um campo complexo e considere a densidade lagrangiana

$$\mathcal{L} = i\hbar\psi^*\dot{\psi} - \frac{\hbar^2}{2m}\nabla\psi\cdot\nabla\psi^* - V(\mathbf{x},t)\psi^*\psi \ . \tag{10.1.27}$$

Tomando ψ e ψ^* como campos independentes, mostre que as equações de Lagrange correspondentes são a equação de Schrödinger

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V(\mathbf{x},t)\psi \tag{10.1.28}$$

e o seu complexo conjugado. ■

10.2 Teorias de Campos Relativísticas

As equações de Lagrange (10.1.21) permanecem inalteradas sob uma mudança de escala das coordenadas $\mathbf{x}, t$. Em particular, fazendo $x^0 = ct$ temos

$$\frac{\partial}{\partial t}\Big(\frac{\partial\mathcal{L}}{\partial(\partial\varphi_\alpha/\partial t)}\Big) = \frac{\partial}{\partial x^0}\Big(\frac{\partial\mathcal{L}}{\partial(\partial\varphi_\alpha/\partial x^0)}\Big) \ . \tag{10.2.1}$$

Suporemos, de ora em diante, que a derivada temporal é sempre em relação a $x^0 = ct$, de modo que, em termos da notação covariante do Capítulo 6, as equações de Lagrange (10.1.21) escrevem-se

$$\partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_\alpha)}\Big) - \frac{\partial\mathcal{L}}{\partial\varphi_\alpha} = 0 \quad , \quad \alpha = 1,\ldots,N \ , \tag{10.2.2}$$

onde

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \Big(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\Big) = \Big(\frac{\partial}{\partial x^0}, \nabla\Big) \ . \tag{10.2.3}$$

Se, para um dado valor de α, φ_α for um campo escalar, $\partial_\mu\varphi_\alpha$ será um quadrivetor covariante. Neste caso, se a lagrangiana[3] $\mathcal{L}$ for uma grandeza escalar, $\partial\mathcal{L}/\partial(\partial_\mu\varphi_\alpha)$ será um quadrivetor contravariante e o primeiro termo à esquerda na Eq.(10.2.2) será um escalar. De modo geral, para que as equações de Lagrange (10.2.2) sejam manifestamente covariantes basta exigir que a lagrangiana seja um escalar. A propósito, uma vez que o elemento de volume quadridimensional d^4x é invariante sob transformações de Lorentz, a ação $S = \int d^4x\mathcal{L}$ também será um escalar se $\mathcal{L}$ for um escalar.

Campo de Klein-Gordon

Como primeiro exemplo de uma teoria de campos relativística, consideremos a teoria de um méson escalar cuja lagrangiana é

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{m^2}{2}\phi^2 \ , \tag{10.2.4}$$

onde m é a massa da partícula (em unidades tais que $\hbar = c = 1$). Esta lagrangiana é

[3]Conscientes do abuso de linguagem, daqui por diante nos referiremos à densidade lagrangiana simplesmente como lagrangiana, como é usual em teoria de campos.

um escalar sob transformações de Lorentz, já que ϕ é um campo escalar (real). A fim de tornar os cálculos mais transparentes, vamos introduzir a notação $\phi_\mu \equiv \partial_\mu \phi$, em termos da qual podemos escrever

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = \frac{\partial}{\partial \phi_\mu}\Big(\frac{1}{2} g^{\nu\lambda}\phi_\nu \phi_\lambda\Big) = \frac{1}{2} g^{\nu\lambda}\delta^\mu_\nu \phi_\lambda + \frac{1}{2} g^{\nu\lambda}\phi_\nu \delta^\mu_\lambda = \frac{1}{2}\big[\, g^{\mu\lambda}\phi_\lambda + g^{\nu\mu}\phi_\nu \,\big] = \phi^\mu \ . \tag{10.2.5}$$

Levando este resultado em (10.2.2) resulta a equação de Klein-Gordon

$$(\Box + m^2)\phi = 0 \ , \tag{10.2.6}$$

onde usamos o operador d'Alembertiano definido pela Eq.(6.4.12). Na teoria quântica de campos esta equação descreve mésons escalares, que são partículas de massa m sem spin.

Campo Eletromagnético

Outro exemplo importante é o do campo eletromagnético na presença de cargas e correntes descritas pela quadricorrente J^μ. Como campos independentes tomemos as componentes do quadripotencial A_μ. Mostremos que a lagrangiana

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}F^{\mu\nu} - \frac{1}{c} J^\mu A_\mu \ , \tag{10.2.7}$$

com

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \ , \tag{10.2.8}$$

dá as equações de campo corretas.[4] Agora $\varphi_\alpha \equiv A_\alpha$ e α é um índice de quadrivetor. Para facilitar o cálculo, definamos o campo tensorial auxiliar

$$A_{\mu\alpha} \equiv \partial_\mu A_\alpha \ . \tag{10.2.9}$$

Tendo em vista que $\partial\mathcal{L}/\partial A_\alpha = -(1/c)J^\alpha$, as equações de Lagrange reduzem-se a

$$\partial_\mu\Big(\frac{\partial \mathcal{L}}{\partial A_{\mu\alpha}}\Big) + \frac{1}{c} J^\alpha = 0 \ . \tag{10.2.10}$$

De (10.2.7) deduz-se

$$\frac{\partial \mathcal{L}}{\partial A_{\mu\alpha}} = -\frac{1}{4}\Big(\frac{\partial F_{\beta\gamma}}{\partial A_{\mu\alpha}} F^{\beta\gamma} + F_{\beta\gamma}\frac{\partial F^{\beta\gamma}}{\partial A_{\mu\alpha}}\Big) = -\frac{1}{2}\frac{\partial F_{\beta\gamma}}{\partial A_{\mu\alpha}} F^{\beta\gamma} \ . \tag{10.2.11}$$

Por outro lado, de $F_{\beta\gamma} = A_{\beta\gamma} - A_{\gamma\beta}$ resulta imediatamente

$$\frac{\partial F_{\beta\gamma}}{\partial A_{\mu\alpha}} = \delta^\mu_\beta \delta^\alpha_\gamma - \delta^\mu_\gamma \delta^\alpha_\beta \ . \tag{10.2.12}$$

[4] A lagrangiana (10.2.7) é adequada no sistema de unidades de Heaviside-Lorentz (Jackson 1999), muito usado na teoria quântica de campos. No sistema gaussiano, o fator 1/4 deve ser substituído por $1/16\pi$.

Conseqüentemente,

$$\frac{\partial \mathcal{L}}{\partial A_{\mu\alpha}} = -\frac{1}{2}(\delta^\mu_\beta \delta^\alpha_\gamma - \delta^\mu_\gamma \delta^\alpha_\beta)F^{\beta\gamma} = -\frac{1}{2}(F^{\mu\alpha} - F^{\alpha\mu}) = F^{\alpha\mu} \tag{10.2.13}$$

e as equações de campo escrevem-se

$$\partial_\mu F^{\mu\alpha} = \frac{1}{c} J^\alpha \ , \tag{10.2.14}$$

que são idênticas a (6.5.13), exceto pela ausência do fator 4π no lado direito porque aqui estamos usando unidades de Heaviside-Lorentz.

As equações de Maxwell homogêneas são automaticamente satisfeitas quando o campo eletromagnético é descrito pelos potenciais ϕ e $\mathbf{A}$. No formalismo manifestamente covariante as equações homogêneas não decorrem da lagrangiana, mas das *identidades*

$$\partial^\alpha F^{\beta\gamma} + \partial^\beta F^{\gamma\alpha} + \partial^\gamma F^{\alpha\beta} = 0 \ . \tag{10.2.15}$$

Estas equações coincidem com as equações de Maxwell homogêneas (vide Seção 6.5) .

■ **Exercício 10.2.1.** Prove que as equações (10.2.15) são identicamente satisfeitas em virtude da definição $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. ■

10.3 Derivadas Funcionais

Na teoria clássica de campos somos levados naturalmente a lidar com funcionais, e a noção de derivada funcional desempenha um papel proeminente. Seja $x = (x_1, \ldots, x_n)$ um ponto de $\boldsymbol{R}^n$ e $f(x)$ uma função suficientemente diferenciável de n variáveis reais. Um funcional F é uma regra que associa um número real $F[f]$ a cada função f. A derivada funcional de F em relação a f no ponto x, denotada por $\delta F/\delta f(x)$, é definida por

$$\left(\frac{d}{d\epsilon} F[f + \epsilon\sigma]\right)_{\epsilon=0} \equiv \int_{\boldsymbol{R}^n} d^n x\, \sigma(x)\, \frac{\delta F}{\delta f(x)} \ . \tag{10.3.1}$$

A partir desta definição implícita prova-se facilmente que a diferenciação funcional satisfaz

$$\frac{\delta}{\delta f(x)}(c_1 F_1 + c_2 F_2) = c_1 \frac{\delta F_1}{\delta f(x)} + c_2 \frac{\delta F_2}{\delta f(x)} \ , \tag{10.3.2}$$

onde c_1 e c_2 são constantes, e obedece à regra de derivada de um produto:

$$\frac{\delta}{\delta f(x)}(F_1 F_2) = \frac{\delta F_1}{\delta f(x)} F_2 + F_1 \frac{\delta F_2}{\delta f(x)} \ . \tag{10.3.3}$$

Prova-se, ainda, que, se $\Phi = \Phi(F)$ é uma função diferenciável do funcional F, vale a regra da cadeia da diferenciação funcional na forma

$$\frac{\delta\Phi}{\delta f(x)} = \frac{d\Phi}{dF}\frac{\delta F}{\delta f(x)} \ . \tag{10.3.4}$$

Vejamos alguns exemplos simples de cálculo de derivadas funcionais. Se o funcional F é definido por

$$F[f] = \int_{R^n} d^n x\, f(x) \ , \tag{10.3.5}$$

então

$$F[f+\epsilon\sigma] = F[f] + \epsilon \int_{R^n} d^n x\, \sigma(x) \implies \Big(\frac{d}{d\epsilon} F[f+\epsilon\sigma]\Big)_{\epsilon=0} = \int_{R^n} d^n x\, \sigma(x) \ , \tag{10.3.6}$$

donde

$$\frac{\delta F}{\delta f(x)} = 1 \ . \tag{10.3.7}$$

Outro exemplo interessante é fornecido por

$$F[f] = \int d^n x\, d^n y\, K(x,y)\, f(x)\, f(y) \ , \tag{10.3.8}$$

onde $K(x,y)$ é uma função conhecida. Temos, agora, omitindo o termo de segunda ordem em ϵ, que não contribui para a derivada funcional,

$$F[f+\epsilon\sigma] = F[f] + \epsilon \int d^n x d^n y K(x,y)\sigma(x) f(y) + \epsilon \int d^n x d^n y K(x,y) f(x)\sigma(y) \ . \tag{10.3.9}$$

Conseqüentemente,

$$\Big(\frac{d}{d\epsilon} F[f+\epsilon\sigma]\Big)_{\epsilon=0} = \int d^n x d^n y K(x,y)\sigma(x) f(y) + \int d^n x d^n y K(x,y) f(x)\sigma(y) \ . \tag{10.3.10}$$

Comparando esta equação com (10.3.1) resulta

$$\frac{\delta F}{\delta f(x)} = \int d^n y \, \big[\, K(x,y) + K(y,x) \,\big] f(y) \ . \tag{10.3.11}$$

Funcionais dependentes de parâmetros ocorrem com freqüência. Um exemplo de interesse é o funcional F_x definido por

$$F_x[f] = \int_{R^n} d^n x'\, K(x,x')\, f(x') \ . \tag{10.3.12}$$

Um cálculo simples e direto mostra que a derivada funcional de F_x é

$$\frac{\delta F_x}{\delta f(y)} = K(x,y) \ . \tag{10.3.13}$$

Com o emprego da "função" delta de Dirac também se pode encarar $f(x)$ como um funcional da forma (10.3.12):

$$f(x) = \int_{R^n} d^n x' \, \delta(x - x') \, f(x') \ . \tag{10.3.14}$$

Temos, de acordo com (10.3.13),

$$\frac{\delta f(x)}{\delta f(y)} = \delta(x - y) \ . \tag{10.3.15}$$

Se $F[f_1, \ldots, f_N]$ é um funcional de N variáveis, a derivada funcional $\delta F / \delta f_k(x)$ é definida por uma extensão natural de (10.3.1), a saber:

$$\Big(\frac{d}{d\epsilon} F[f_1 + \epsilon\sigma_1, \ldots, f_N + \epsilon\sigma_N]\Big)_{\epsilon=0} \equiv \int d^n x \sum_{k=1}^{N} \sigma_k(x) \frac{\delta F}{\delta f_k(x)} \ . \tag{10.3.16}$$

Comparando esta definição de derivada funcional com a noção de variação de um funcional introduzida na Seção 2.2, podemos escrever

$$\delta F = \int d^n x \sum_{k=1}^{N} \frac{\delta F}{\delta f_k(x)} \, \delta f_k(x) \ . \tag{10.3.17}$$

Em outras palavras, a derivada funcional de F é determinada pelo termo de primeira ordem da expansão de $F[f + \delta f]$ em potências de δf. A Eq.(10.3.15) adquire a forma generalizada

$$\frac{\delta f_i(x)}{\delta f_j(y)} = \delta_{ij} \, \delta(x - y) \ . \tag{10.3.18}$$

Inspecionando (10.1.26) e comparando com (10.3.17), somos conduzidos a

$$\frac{\delta S}{\delta \varphi_\alpha(x)} = \frac{\partial \mathcal{L}}{\partial \varphi_\alpha(x)} - \frac{\partial}{\partial t}\Big(\frac{\partial \mathcal{L}}{\partial \dot{\varphi}_\alpha(x)}\Big) - \boldsymbol{\nabla} \cdot \Big(\frac{\partial \mathcal{L}}{\partial (\boldsymbol{\nabla} \varphi_\alpha(x))}\Big) \ , \tag{10.3.19}$$

onde $x = (\mathbf{x}, t)$. Portanto, em termos de derivadas funcionais da ação, as equações de Lagrange podem ser escritas na forma compacta

$$\frac{\delta S}{\delta \varphi_\alpha(x)} = 0 \ . \tag{10.3.20}$$

As derivadas funcionais são particularmente úteis na formulação hamiltoniana das teorias de campos, que passamos a discutir.

10.4 Teoria de Campos na Forma Hamiltoniana

Vamos definir o momento canonicamente conjugado a $\varphi_\alpha(x)$, denotado por $\pi^\alpha(x)$, da mesma forma que na dinâmica de partículas:

$$\pi^\alpha(x) = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}_\alpha(x)} \ . \tag{10.4.1}$$

Suponhamos que as Eqs.(10.4.1) sejam solúveis para os $\dot{\varphi}_\alpha$. Neste caso, a densidade hamiltoniana $\mathcal{H}$, interpretada como a densidade de energia e definida por

$$\mathcal{H} = \sum_\alpha \pi^\alpha \, \dot{\varphi}_\alpha - \mathcal{L} \ , \tag{10.4.2}$$

pode ser expressa em termos de π^α , φ_α e seus gradientes. A hamiltoniana

$$H[\varphi_\alpha, \pi^\alpha] = \int d^3x \, \mathcal{H}\big(\varphi_\alpha(x), \boldsymbol{\nabla}\varphi_\alpha(x), \pi^\alpha(x), \boldsymbol{\nabla}\pi^\alpha(x)\big) \tag{10.4.3}$$

é um funcional dos campos e de seus momentos conjugados.[5]

Equações de Hamilton

A ação na forma hamiltoniana escreve-se

$$S = \int_\Omega d^4x \left\{ \sum_\alpha \pi^\alpha \, \dot{\varphi}_\alpha \ - \ \mathcal{H}\big(\varphi_\alpha, \boldsymbol{\nabla}\varphi_\alpha, \pi^\alpha, \boldsymbol{\nabla}\pi^\alpha\big) \right\} \tag{10.4.4}$$

e as equações de Hamilton decorrem do princípio variacional $\delta S = 0$. Variando independentemente os campos e seus momentos canonicamente conjugados, o princípio de Hamilton toma a forma

$$\delta S = \int_\Omega d^4x \sum_\alpha \left\{ \pi^\alpha \delta\dot{\varphi}_\alpha + \delta\pi^\alpha \dot{\varphi}_\alpha - \frac{\partial \mathcal{H}}{\partial \varphi_\alpha}\delta\varphi_\alpha - \frac{\partial \mathcal{H}}{\partial(\boldsymbol{\nabla}\varphi_\alpha)} \cdot \delta(\boldsymbol{\nabla}\varphi_\alpha) - \frac{\partial \mathcal{H}}{\partial \pi^\alpha}\delta\pi^\alpha - \frac{\partial \mathcal{H}}{\partial(\boldsymbol{\nabla}\pi^\alpha)} \cdot \delta(\boldsymbol{\nabla}\pi^\alpha) \right\}$$

$$= \int_\Omega d^4x \sum_\alpha \left\{ \Big(-\dot{\pi}^\alpha - \frac{\partial \mathcal{H}}{\partial \varphi_\alpha} + \boldsymbol{\nabla} \cdot \frac{\partial \mathcal{H}}{\partial(\boldsymbol{\nabla}\varphi_\alpha)}\Big)\delta\varphi_\alpha + \Big(\dot{\varphi}_\alpha - \frac{\partial \mathcal{H}}{\partial \pi^\alpha} + \boldsymbol{\nabla} \cdot \frac{\partial \mathcal{H}}{\partial(\boldsymbol{\nabla}\pi^\alpha)}\Big)\delta\pi^\alpha \right\} = 0 \ , \tag{10.4.5}$$

onde, como de hábito, uma integração por partes foi feita e termos de fronteira foram descartados. Igualando a zero os coeficientes de $\delta\varphi_\alpha$ e $\delta\pi^\alpha$ obtemos

$$\dot{\varphi}_\alpha = \frac{\partial \mathcal{H}}{\partial \pi^\alpha} - \boldsymbol{\nabla} \cdot \frac{\partial \mathcal{H}}{\partial(\boldsymbol{\nabla}\pi^\alpha)} \ , \tag{10.4.6a}$$

$$\dot{\pi}^\alpha = -\frac{\partial \mathcal{H}}{\partial \varphi_\alpha} + \boldsymbol{\nabla} \cdot \frac{\partial \mathcal{H}}{\partial(\boldsymbol{\nabla}\varphi_\alpha)} \ , \tag{10.4.6b}$$

[5] Em todas as teorias de interesse físico fundamental, $\mathcal{H}$ não depende dos gradientes dos π^α.

que são as equações de campo na forma hamiltoniana. Usando (10.3.19) com H no lugar de S e notando que $\mathcal{H}$ não depende das derivadas temporais dos φ_α e π^α, podemos escrever as equações de Hamilton (10.4.6) na forma condensada

$$\dot{\varphi}_\alpha(x) = \frac{\delta H}{\delta \pi^\alpha(x)} \quad , \quad \dot{\pi}^\alpha(x) = -\frac{\delta H}{\delta \varphi_\alpha(x)} \quad , \tag{10.4.7}$$

cuja semelhança com as equações de Hamilton para sistemas discretos é notável.

■ **Exemplo 10.4.1.** Obter a densidade hamiltoniana e as equações de Hamilton para o campo de Klein-Gordon descrito pela lagrangiana (10.2.4).

Solução. Escrevendo a lagrangiana (10.2.4) na forma

$$\mathcal{L} = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}\nabla\phi\cdot\nabla\phi - \frac{m^2}{2}\phi^2 \tag{10.4.8}$$

resulta

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(x)} = \dot{\phi}(x) \ . \tag{10.4.9}$$

Portanto,

$$\mathcal{H} = \pi\dot{\phi} - \mathcal{L} = \frac{1}{2}\pi^2 + \frac{1}{2}\nabla\phi\cdot\nabla\phi + \frac{m^2}{2}\phi^2 \tag{10.4.10}$$

e as equações de Hamilton (10.4.6) adquirem a forma

$$\dot{\phi}(x) = \pi(x) \quad , \quad \dot{\pi}(x) = \nabla^2\phi(x) - m^2\phi(x) \ . \tag{10.4.11}$$

É fácil verificar que estas equações são equivalentes à equação de Klein-Gordon (10.2.6). ■

Parênteses de Poisson

Na formulação hamiltoniana, uma variável dinâmica X é um funcional dos φ_α e π^α da forma

$$X[\varphi,\pi] = \int d^3x\, \mathcal{X}\big(\varphi_\alpha, \nabla\varphi_\alpha, \pi^\alpha, \nabla\pi^\alpha; \mathbf{x}, t\big) \ , \tag{10.4.12}$$

onde a integração estende-se por todo o espaço e supomos que $\varphi_\alpha(x)$ e $\pi^\alpha(x)$ anulam-se no infinito espacial. Note que as variáveis dinâmicas são, geralmente, funções do tempo. O parêntese de Poisson de duas variáveis dinâmicas é definido por

$$\{X,Y\} = \int d^3x \sum_{\alpha=1}^{N} \left(\frac{\delta X}{\delta\varphi_\alpha(x)}\frac{\delta Y}{\delta\pi^\alpha(x)} - \frac{\delta X}{\delta\pi^\alpha(x)}\frac{\delta Y}{\delta\varphi_\alpha(x)} \right) \ . \tag{10.4.13}$$

Com o intuito de determinar a equação de movimento de uma variável dinâmica arbitrária, note que a variação de X num intervalo de tempo δt é dada por

$$\delta X = \int d^3x \sum_{\alpha=1}^{N} \Big\{ \frac{\delta X}{\delta\varphi_\alpha(x)} \dot{\varphi}_\alpha(x)\delta t + \frac{\delta X}{\delta\pi^\alpha(x)} \dot{\pi}^\alpha(x)\delta t + \frac{\partial \mathcal{X}}{\partial t}\delta t \Big\} \ , \tag{10.4.14}$$

uma vez que $\delta\varphi_\alpha = \dot{\varphi}_\alpha \delta t$ e $\delta\pi^\alpha = \dot{\pi}^\alpha \delta t$. Usando as equações de Hamilton (10.4.7) e a definição de parêntese de Poisson (10.4.13), resulta

$$\frac{dX}{dt} = \{X, H\} + \frac{\partial X}{\partial t} \ , \tag{10.4.15}$$

onde, *por definição*,

$$\frac{\partial X}{\partial t} = \int d^3x \, \frac{\partial \mathcal{X}}{\partial t} \ . \tag{10.4.16}$$

Uma das principais virtudes da notação funcional é tornar a equação de movimento para uma variável dinâmica na teoria de campos formalmente idêntica à Eq.(8.4.3) para sistemas discretos.

Considere os funcionais

$$X_{\mathbf{x}}[\varphi_\alpha] = \varphi_\alpha(\mathbf{x}, t) = \int d^3x' \, \delta(\mathbf{x} - \mathbf{x}') \, \varphi_\alpha(\mathbf{x}', t) \tag{10.4.17}$$

e

$$Y_{\mathbf{x}}[\pi^\beta] = \pi^\beta(\mathbf{x}, t) = \int d^3x' \, \delta(\mathbf{x} - \mathbf{x}') \, \pi^\beta(\mathbf{x}', t) \ . \tag{10.4.18}$$

Com o emprego de (10.3.18) obtém-se

$$\{X_{\mathbf{x}}, Y_{\mathbf{y}}\} = \int d^3x' \sum_{\gamma=1}^{N} \left\{ \frac{\delta\varphi_\alpha(\mathbf{x}, t)}{\delta\varphi_\gamma(\mathbf{x}', t)} \frac{\delta\pi^\beta(\mathbf{y}, t)}{\delta\pi^\gamma(\mathbf{x}', t)} - \frac{\delta\varphi_\alpha(\mathbf{x}, t)}{\delta\pi^\gamma(\mathbf{x}', t)} \frac{\delta\pi^\beta(\mathbf{y}, t)}{\delta\varphi_\gamma(\mathbf{x}', t)} \right\}$$

$$= \int d^3x' \sum_{\gamma=1}^{N} \delta_\alpha^\gamma \, \delta(\mathbf{x} - \mathbf{x}') \, \delta_\gamma^\beta \, \delta(\mathbf{y} - \mathbf{x}') = \delta_\alpha^\beta \, \delta(\mathbf{x} - \mathbf{y}) \ . \tag{10.4.19}$$

Este resultado costuma ser expresso na forma

$$\{\varphi_\alpha(\mathbf{x}, t), \pi^\beta(\mathbf{y}, t)\} = \delta_\alpha^\beta \, \delta(\mathbf{x} - \mathbf{y}) \ , \tag{10.4.20}$$

que é a versão contínua de $\{q_i, p_j\} = \delta_{ij}$. Analogamente, é imediato que

$$\{\varphi_\alpha(\mathbf{x}, t), \varphi_\beta(\mathbf{y}, t)\} = 0 \quad , \quad \{\pi^\alpha(\mathbf{x}, t), \pi^\beta(\mathbf{y}, t)\} = 0 \ . \tag{10.4.21}$$

Os parênteses de Poisson fundamentais (10.4.20) e (10.4.21) servem de ponto de partida para a quantização canônica de uma teoria clássica de campos. O parêntese de Poisson dos campos clássicos multiplicado por $i\hbar$ é substituído pelo comutador dos operadores de campo correspondentes, no caso de campos bosônicos, ou pelo anticomutador, no caso de campos fermiônicos.

10.5 Simetrias da Ação e Teorema de Noether

A versão contínua do teorema de Noether estabelece a conexão genérica entre simetrias e leis de conservação na teoria clássica de campos.

Transformação Infinitesimal

Consideremos a transformação infinitesimal

$$x^\mu \longrightarrow x'^\mu = x^\mu + \Delta x^\mu \ , \tag{10.5.1}$$

$$\varphi_\alpha(x) \longrightarrow \varphi'_\alpha(x') = \varphi_\alpha(x) + \Delta\varphi_\alpha(x) \ , \tag{10.5.2}$$

com $\Delta x^\mu = \Delta x^\mu(x)$. A variação Δ difere da variação usual δ porque leva em conta a mudança do campo em conseqüência da alteração tanto de sua forma funcional quanto de seu argumento. A variação provocada apenas pela mudança de forma é definida por

$$\delta\varphi_\alpha(x) = \varphi'_\alpha(x) - \varphi_\alpha(x) \ , \tag{10.5.3}$$

de modo que

$$\Delta\varphi_\alpha(x) = \varphi'_\alpha(x') - \varphi_\alpha(x') + \varphi_\alpha(x') - \varphi_\alpha(x) = \delta\varphi_\alpha(x') + \partial_\mu\varphi_\alpha(x)\,\Delta x^\mu \ , \tag{10.5.4}$$

onde estamos empregando a convenção de soma sobre índices repetidos e a notação covariante do Capítulo 6. Desprezando termos de segunda ordem nas variações infinitesimais, a Eq.(10.5.4) se reduz a

$$\Delta\varphi_\alpha(x) = \delta\varphi_\alpha(x) + \varphi_{\alpha;\mu}\,\Delta x^\mu \ , \tag{10.5.5}$$

onde introduzimos a notação abreviada $\varphi_{\alpha;\mu}$ para as derivadas dos campos:

$$\varphi_{\alpha;\mu}(x) \equiv \partial_\mu\,\varphi_\alpha(x) \ . \tag{10.5.6}$$

É importante destacar que as derivadas ∂_μ comutam com a operação δ mas *não comutam* com a operação Δ, isto é, $\Delta(\varphi_{\alpha;\beta}) \neq (\Delta\varphi_\alpha)_{;\beta}$. Aplicando a Eq.(10.5.5) a $\varphi_{\alpha;\beta}$ resulta

$$\Delta\varphi_{\alpha;\beta}(x) = \delta\varphi_{\alpha;\beta}(x) + \varphi_{\alpha;\beta;\mu}\,\Delta x^\mu \ . \tag{10.5.7}$$

Invariância da Ação

A variação da ação é definida por

$$\Delta S = \int_{\Omega'} d^4x'\mathcal{L}\big(\varphi'_\alpha(x'), \varphi'_{\alpha;\beta}(x'), x'\big) - \int_\Omega d^4x\,\mathcal{L}\big(\varphi_\alpha(x), \varphi_{\alpha;\beta}(x), x\big) \ . \tag{10.5.8}$$

A lagrangiana variada é dada por

$$\begin{aligned}\mathcal{L}' &= \mathcal{L}\big(\varphi'_\alpha(x'), \varphi'_{\alpha;\beta}(x'), x'\big) = \mathcal{L}\big(\varphi_\alpha(x) + \Delta\varphi_\alpha(x), \varphi_{\alpha;\beta}(x) + \Delta\varphi_{\alpha;\beta}(x), x + \Delta x\big) \\ &= \mathcal{L}\big(\varphi_\alpha(x), \varphi_{\alpha;\beta}(x), x\big) + \frac{\partial\mathcal{L}}{\partial\varphi_\alpha}\Delta\varphi_\alpha + \frac{\partial\mathcal{L}}{\partial\varphi_{\alpha;\beta}}\Delta\varphi_{\alpha;\beta} + \frac{\partial\mathcal{L}}{\partial x^\mu}\Delta x^\mu \\ &= \mathcal{L} + \frac{\partial\mathcal{L}}{\partial\varphi_\alpha}\delta\varphi_\alpha + \frac{\partial\mathcal{L}}{\partial\varphi_{\alpha;\beta}}\delta\varphi_{\alpha;\beta} + \frac{d\mathcal{L}}{dx^\mu}\Delta x^\mu = \mathcal{L} + \delta\mathcal{L} + \frac{d\mathcal{L}}{dx^\mu}\Delta x^\mu \ , \qquad (10.5.9)\end{aligned}$$

onde usamos (10.5.5) e (10.5.7), e d/dx^μ denota o que é lícito chamar de "derivada parcial total" em relação à coordenada x^μ,

$$\frac{d\mathcal{L}}{dx^\mu} = \frac{\partial\mathcal{L}}{\partial\varphi_\alpha}\frac{\partial\varphi_\alpha}{\partial x^\mu} + \frac{\partial\mathcal{L}}{\partial\varphi_{\alpha;\beta}}\frac{\partial\varphi_{\alpha;\beta}}{\partial x^\mu} + \frac{\partial\mathcal{L}}{\partial x^\mu} \ , \qquad (10.5.10)$$

que leva em conta a dependência explícita e a dependência por intermédio dos campos.

Por outro lado, pelo item (i) do Problema 8.19, temos

$$d^4x' = \frac{\partial(x'^0, x'^1, x'^2, x'^3)}{\partial(x^0, x^1, x^2, x^3)}\, d^4x = \Big(1 + \frac{\partial\Delta x^\mu}{\partial x^\mu}\Big)\, d^4x \ . \qquad (10.5.11)$$

Levando (10.5.9) e (10.5.11) em (10.5.8), e desprezando termos de ordem superior à primeira nas variações, resulta

$$\Delta S = \int_\Omega d^4x \Big[\delta\mathcal{L} + \frac{d\mathcal{L}}{dx^\mu}\Delta x^\mu + \mathcal{L}\,\frac{\partial\Delta x^\mu}{\partial x^\mu}\Big] = \int_\Omega d^4x \Big[\delta\mathcal{L} + \frac{d}{dx^\mu}\big(\mathcal{L}\,\Delta x^\mu\big)\Big] \ . \qquad (10.5.12)$$

Usando as equações de movimento

$$\frac{\partial\mathcal{L}}{\partial\varphi_\alpha} = \frac{d}{dx^\mu}\Big(\frac{\partial\mathcal{L}}{\partial\varphi_{\alpha;\mu}}\Big) \qquad (10.5.13)$$

obtemos

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\varphi_\alpha}\delta\varphi_\alpha + \frac{\partial\mathcal{L}}{\partial\varphi_{\alpha;\beta}}\delta\varphi_{\alpha;\beta} = \frac{d}{dx^\mu}\Big(\frac{\partial\mathcal{L}}{\partial\varphi_{\alpha;\mu}}\Big)\delta\varphi_\alpha + \frac{\partial\mathcal{L}}{\partial\varphi_{\alpha;\beta}}\frac{\partial}{\partial x^\beta}\big(\delta\varphi_\alpha\big) = \frac{d}{dx^\mu}\Big(\frac{\partial\mathcal{L}}{\partial\varphi_{\alpha;\mu}}\,\delta\varphi_\alpha\Big) \ , \qquad (10.5.14)$$

onde, para obter a última igualdade, o índice mudo de soma β foi trocado por μ. Introduzindo (10.5.14) em (10.5.12), a condição de invariância da ação torna-se

$$\Delta S = \int_\Omega d^4x\, \frac{d}{dx^\mu}\left\{\frac{\partial\mathcal{L}}{\partial\varphi_{\alpha;\mu}}\,\delta\varphi_\alpha + \mathcal{L}\,\Delta x^\mu\right\} = 0 \ , \qquad (10.5.15)$$

que implica uma lei de conservação local na forma covariante (6.5.6). De fato, a natureza arbitrária do domínio de integração exige que o integrando na Eq.(10.5.15) seja zero, o que equivale à equação da continuidade para a quadricorrente definida pela expressão entre chaves.

Quantidades Conservadas

É mais conveniente, no entanto, exprimir a lei de conservação da quadricorrente em termos dos parâmetros infinitesimais que caracterizam a transformação. Suponha que a transformação (10.5.1-2) seja especificada em termos de R parâmetros infinitesimais independentes $\epsilon_1, \ldots, \epsilon_R$ na forma

$$\Delta x^\mu = \sum_{r=1}^{R} X^{\mu\,(r)} \epsilon_r \quad , \qquad \Delta\varphi_\alpha = \sum_{r=1}^{R} \Psi_\alpha^{(r)}\, \epsilon_r \ . \tag{10.5.16}$$

Os índices α e r dos campos e dos parâmetros da transformação podem ou não ter caráter tensorial, e persistiremos convencionando que índices repetidos de qualquer natureza indicam uma soma. Substituindo (10.5.16) em (10.5.5) podemos escrever

$$\Delta x^\mu = X^{\mu\,(r)} \epsilon_r \quad , \qquad \delta\varphi_\alpha = \left(\Psi_\alpha^{(r)} - \varphi_{\alpha;\nu} X^{\nu\,(r)} \right) \epsilon_r \ , \tag{10.5.17}$$

que, introduzidas em (10.5.15), conduzem a

$$\Delta S = -\int_\Omega d^4x\, \epsilon_r\, \partial_\mu\, \Theta^{\mu\,(r)} = 0 \ , \tag{10.5.18}$$

onde

$$\Theta^{\mu\,(r)} = -\frac{\partial \mathcal{L}}{\partial \varphi_{\alpha;\mu}} \left(\Psi_\alpha^{(r)} - \varphi_{\alpha;\nu} X^{\nu\,(r)} \right) - \mathcal{L}\, X^{\mu\,(r)} \ . \tag{10.5.19}$$

Como a região de integração e os parâmetros ϵ_r são arbitrários, de (10.5.18) deduzem-se as R leis de conservação local[6]

$$\partial_\mu\, \Theta^{\mu\,(r)} = 0 \quad , \quad r = 1, \ldots, R \ . \tag{10.5.20}$$

Fazendo $\Theta^{\mu\,(r)} = (\Theta^{0\,(r)}, \boldsymbol{\Theta}^{(r)})$ podemos escrever (10.5.20) na forma usual

$$\partial_0\, \Theta^{0\,(r)} + \boldsymbol{\nabla} \cdot \boldsymbol{\Theta}^{(r)} = 0 \ , \tag{10.5.21}$$

donde

$$\frac{d}{dx^0} \int_V d^3x\, \Theta^{0\,(r)} = \int_V d^3x\, \partial_0\, \Theta^{0\,(r)} = -\int_V d^3x\, \boldsymbol{\nabla} \cdot \boldsymbol{\Theta}^{(r)} = -\oint_\Sigma \boldsymbol{\Theta}^{(r)} \cdot \mathbf{da} \ , \tag{10.5.22}$$

onde Σ é a superfície bidimensional que limita o volume V e usamos o teorema da divergência. Se a região de integração V engloba todo o espaço tridimensional e os campos tendem a zero com suficiente rapidez no infinito, a integral de superfície é nula e as integrais

[6] Não havendo mais risco de ambigüidade, a partir da Eq.(10.5.18) passamos a denotar pelo símbolo ∂_μ a derivada parcial total em relação a x^μ.

$$C^{(r)} = \int d^3x\, \Theta^{0\,(r)} \tag{10.5.23}$$

são quantidades conservadas, pois não dependem do tempo. Fica, assim, demonstrado que a cada transformação infinitesimal R-paramétrica das coordenadas e campos que deixa a ação invariante correspondem R quantidades conservadas $C^{(r)}$ dadas por (10.5.23), conhecidas como *cargas de Noether*.

Conservação do Quadrimomento

Como primeiro exemplo importante, considere a "translação" no espaço-tempo

$$x'^{\mu} = x^{\mu} + \epsilon^{\mu} \ , \tag{10.5.24}$$

que não modifica os campos, isto é, $\Delta\varphi_\alpha = 0$. Como o jacobiano desta transformação é 1, as Eqs.(10.5.8) e (10.5.9) mostram que a ação é invariante desde que a lagrangiana não dependa explicitamente das coordenadas do espaço-tempo. O índice r dos parâmetros da transformação tem natureza quadrivetorial e, de acordo com (10.5.16), temos

$$X^{\mu(\beta)} = g^{\mu\beta} \quad , \quad \Psi_{\alpha}^{\ (\beta)} = 0 \ . \tag{10.5.25}$$

Substituindo estas expressões em (10.5.19) obtemos o *tensor de energia-momento*

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha;\mu}} \varphi_{\alpha;\beta}\, g^{\beta\nu} - \mathcal{L}\, g^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha;\mu}} \frac{\partial \varphi_\alpha}{\partial x_\nu} - \mathcal{L}\, g^{\mu\nu} \ . \tag{10.5.26}$$

As cargas de Noether conservadas compõem o quadrivetor

$$P^{\nu} = \int d^3x\, T^{0\nu} \ . \tag{10.5.27}$$

A componente zero deste quadrivetor é a integral da densidade hamiltoniana, ou seja, é a energia. Considerações de covariância – e o fato de a conservação do momento linear estar associada à invariância sob translações espaciais – estabelecem que P^ν é o quadrimomento ou quadrivetor de energia-momento, daí chamar-se $T^{\mu\nu}$ de tensor de energia-momento.

■ **Exercício 10.5.1.** Mostre que o tensor de energia-momento do campo escalar descrito pela lagrangiana de Klein-Gordon (10.2.4) é

$$T^{\mu\nu} = \partial^\mu\phi\, \partial^\nu\phi - \frac{1}{2}\left(\partial_\alpha\phi\, \partial^\alpha\phi - m^2\phi^2\right) g^{\mu\nu} \tag{10.5.28}$$

e obtenha a expressão para o momento linear (conservado) do campo. ■

Transformação de Calibre e Conservação da Carga Elétrica

Considere, agora, um campo escalar complexo descrito pela lagrangiana

$$\mathcal{L} = \partial_\mu \phi^* \, \partial^\mu \phi - m^2 \phi^* \phi \ , \tag{10.5.29}$$

a qual é obviamente invariante sob a transformação uniparamétrica

$$\phi' = e^{i\lambda} \, \phi \quad , \quad \phi^{*\prime} = e^{-i\lambda} \, \phi^* \ , \tag{10.5.30}$$

onde λ é um número real arbitrário. Esta transformação é conhecida como *transformação de calibre global* ou *transformação de calibre de primeira espécie*, e sua versão infinitesimal com $\lambda = \epsilon$ escreve-se

$$\Delta\phi = i\epsilon \, \phi \quad , \quad \Delta\phi^* = -i\epsilon \, \phi^* \ , \tag{10.5.31}$$

donde

$$X^{\mu\,(1)} = 0 \quad , \quad \Psi_1^{(1)} = i \quad , \quad \Psi_2^{(1)} = -i \ , \tag{10.5.32}$$

com $\varphi_1 = \phi$, $\varphi_2 = \phi^*$ e o índice r assumindo apenas o valor $r = 1$. Como não há mudança nas coordenadas, a invariância da lagrangiana implica a invariância da ação. A corrente de Noether (10.5.19) toma a forma

$$J^\mu = -\frac{\partial \mathcal{L}}{\partial \varphi_{1;\mu}} \Psi_1^{(1)} - \frac{\partial \mathcal{L}}{\partial \varphi_{2;\mu}} \Psi_2^{(1)} = -i \left(\phi^* \, \partial^\mu \phi - \phi \, \partial^\mu \phi^* \right) \ . \tag{10.5.33}$$

A carga de Noether conservada, interpretada como carga elétrica, é

$$Q = i \int d^3x \left(\phi \, \dot{\phi}^* - \dot{\phi} \, \phi^* \right) \ . \tag{10.5.34}$$

Um campo escalar real não pode interagir com um campo eletromagnético porque a densidade de corrente J^μ é zero (mais detalhes no Problema 10.10). Portanto, campos escalares complexos são necessários para descrever partículas carregadas sem spin.

■ **Exercício 10.5.2.** Prove que a lagrangiana (10.1.27) do campo de Schrödinger é invariante sob transformações de calibre de primeira espécie e encontre a corrente de Noether e a carga de Noether conservada (probabilidade de localização da partícula, na mecânica quântica). ■

10.6 Ondas Solitárias e Sólitons

Configurações de campo com energia finita e localizada que se deslocam sem mudança de forma nem diminuição da velocidade são chamadas de *ondas solitárias*. Essas ondas aparecem tipicamente em teorias de campos não-lineares (Lee 1981) e foram discutidas originalmente por J. Scott Russell (1844), que assim relatou seu primeiro contato com o fenômeno, ocorrido em agosto de 1834:

> Eu estava observando o movimento de um barco, que era puxado rapidamente ao longo de um canal estreito por um par de cavalos, quando o barco parou de repente – mas não a massa de água no canal que ele havia posto em movimento; esta acumulou-se ao redor da proa do barco num estado de violenta agitação e, em seguida, deixando-o repentinamente para trás, rolou para a frente a grande velocidade, assumindo a forma de uma grande elevação solitária, um monte de água arredondado, suave e bem definido que continuou seu curso ao longo do canal aparentemente sem mudança de forma ou diminuição de velocidade. Eu a segui a cavalo e a alcancei ainda rolando a uma velocidade de umas oito ou nove milhas por hora, preservando seu contorno original de uns trinta pés de comprimento por um pé a um pé e meio de altura. Sua altura diminuiu gradualmente e, depois de uma perseguição por uma ou duas milhas, eu a perdi nos meandros do canal.

Como exemplo de modelo que admite soluções do tipo descrito acima, considere um campo escalar real φ em uma dimensão espacial com lagrangiana

$$\mathcal{L} = \frac{1}{2}\,\partial^\mu\varphi\partial_\mu\varphi - \frac{\lambda^2}{8}\,(\varphi^2 - a^2)^2 = \frac{1}{2}\left[\left(\frac{\partial\varphi}{\partial t}\right)^2 - \left(\frac{\partial\varphi}{\partial x}\right)^2\right] - \frac{\lambda^2}{8}\,(\varphi^2 - a^2)^2 \ , \qquad (10.6.1)$$

onde λ e a são constantes positivas. A equação de Lagrange para φ é

$$\frac{\partial^2\varphi}{\partial t^2} - \frac{\partial^2\varphi}{\partial x^2} + \frac{\lambda^2}{2}\varphi(\varphi^2 - a^2) = 0 \ . \qquad (10.6.2)$$

Usando $\pi = \partial\mathcal{L}/\partial\dot\varphi = \dot\varphi$, encontra-se a densidade hamiltoniana

$$\mathcal{H} = \pi\dot\varphi - \mathcal{L} = \frac{1}{2}\left[\left(\frac{\partial\varphi}{\partial t}\right)^2 + \left(\frac{\partial\varphi}{\partial x}\right)^2\right] + \frac{\lambda^2}{8}\,(\varphi^2 - a^2)^2 \ . \qquad (10.6.3)$$

Note que $\varphi = a$ ou $\varphi = -a$ são soluções estáticas da equação de movimento (10.6.2) com energia zero (porque $\mathcal{H} = 0$), que é a menor energia possível. Assim, essa teoria possui duas configurações de campo distintas de energia mínima, isto é, seu estado fundamental[7] é degenerado.

Busquemos outras soluções estáticas $\varphi_0(x)$ que se reduzam assintoticamente ao estado fundamental. Devemos ter

$$-\varphi_0'' + \frac{\lambda^2}{2}\varphi_0(\varphi_0^2 - a^2) = 0 \ . \qquad (10.6.4)$$

Multiplicando esta equação por φ_0' deduz-se imediatamente

$$\frac{1}{2}\frac{d}{dx}\left[(\varphi_0')^2 - \frac{\lambda^2}{4}(\varphi_0^2 - a^2)^2\right] = 0 \implies (\varphi_0')^2 = \frac{\lambda^2}{4}(\varphi_0^2 - a^2)^2 + C \ , \qquad (10.6.5)$$

onde C é uma constante. A condição $\varphi_0(x) \to \pm a$ para $x \to \pm\infty$ requer $C = 0$, de modo que

$$(\varphi_0')^2 = \frac{\lambda^2}{4}(\varphi_0^2 - a^2)^2 \ . \qquad (10.6.6)$$

[7] Também chamado de "vácuo".

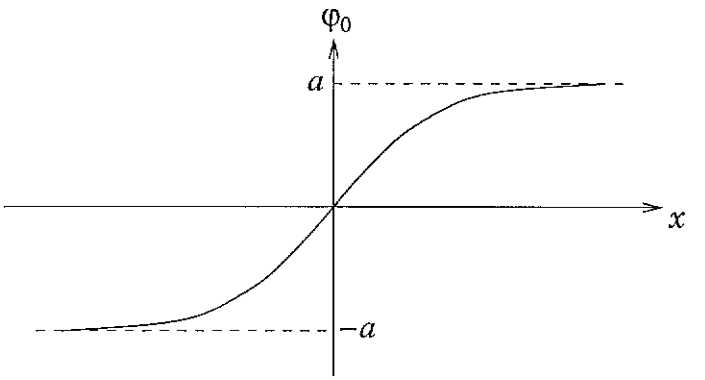


Figura 10.3: O "kink" (10.6.8) conecta dois estados fundamentais ("vácuos") distintos em $x = \pm\infty$.

Extraindo a raiz quadrada negativa ficamos com

$$\frac{d\varphi_0}{dx} = -\frac{\lambda}{2}(\varphi_0^2 - a^2) \ . \tag{10.6.7}$$

Esta equação é imediatamente solúvel por separação de variáveis, dando[8]

$$\varphi_0(x;\lambda) = a\tanh\Big(\frac{\lambda a x}{2}\Big) \ . \tag{10.6.8}$$

Esta configuração de campo (ver Fig. 10.3) conecta o estado fundamental $\varphi = -a$ em $x = -\infty$ ao estado fundamental $\varphi = a$ em $x = \infty$, e costuma ser chamada de um "kink".[9] A densidade de energia correspondente é

$$\mathcal{H}_0 = \frac{1}{2}\Big(\frac{\partial\varphi_0}{\partial x}\Big)^2 + \frac{\lambda^2}{8}(\varphi_0^2 - a^2)^2 = \frac{\lambda^2}{4}(\varphi_0^2 - a^2)^2 = \frac{1}{4}\,\frac{\lambda^2 a^4}{\cosh^4(\lambda a x/2)} \ , \tag{10.6.9}$$

onde usamos (10.6.6) e a identidade $\operatorname{sech}^2 u = \tanh^2 u - 1$. A Fig. 10.4 mostra que trata-se de uma configuração de campo com energia localizada em torno de $x = 0$. Portanto, a energia é finita com valor

$$E_0 = \int_{-\infty}^{\infty}\mathcal{H}_0 dx = \frac{\lambda a^3}{2}\int_{-\infty}^{\infty}\frac{dy}{\cosh^4 y} = \frac{\lambda a^3}{2}\left[\frac{\operatorname{senh} y}{3\cosh^3 y} + \frac{2}{3}\tanh y\right]_{-\infty}^{\infty} = \frac{2\lambda a^3}{3} \ . \tag{10.6.10}$$

A invariância relativística da teoria sugere que se fizermos a substituição $x \longrightarrow \gamma(x - vt)$

[8] A solução geral é da forma (10.6.8) com $x - x_0$ no lugar de x; tomamos $x_0 = 0$ para simplificar.
[9] Em inglês, dobra ou torção. A outra solução possível $\varphi = -a\tanh(\lambda a x/2)$ é chamada de "antikink".

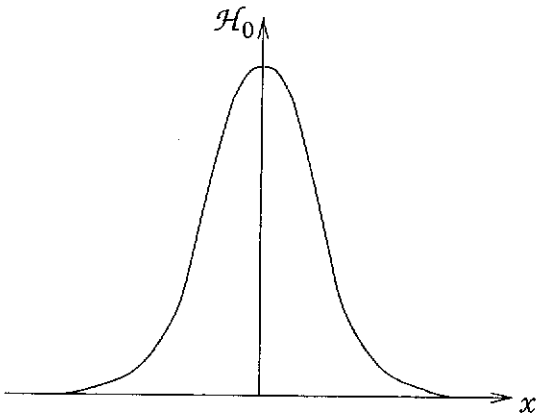


Figura 10.4: Densidade de energia como função da posição do "kink" (10.6.8).

com $\gamma = (1-v^2)^{-1/2}$ obteremos uma solução viajante com velocidade constante v para o campo (estamos usando unidades tais que $c = 1$). Seja, portanto,

$$\varphi(x,t) = a \tanh\left[\frac{a\lambda}{2}\,\frac{x-vt}{\sqrt{1-v^2}}\right] , \tag{10.6.11}$$

que, pela Eq.(10.6.8), pode ser reescrita na forma

$$\varphi(x,t) = \varphi_0(\xi;\lambda') \quad , \quad \xi = x - vt \quad , \quad \lambda' = \frac{\lambda}{\sqrt{1-v^2}} . \tag{10.6.12}$$

Notando que

$$\ddot{\varphi}(x,t) = v^2\varphi_0''(\xi;\lambda') \quad , \quad \varphi''(x,t) = \varphi_0''(\xi;\lambda') \tag{10.6.13}$$

e usando a Eq.(10.6.4) com λ' no lugar de λ, temos

$$\frac{\partial^2\varphi}{\partial t^2} - \frac{\partial^2\varphi}{\partial x^2} + \frac{\lambda^2}{2}\varphi(\varphi^2-a^2) = (v^2-1)\varphi_0''(\xi;\lambda') + \frac{\lambda^2}{2}\varphi_0(\xi;\lambda')\big(\varphi_0(\xi;\lambda')^2 - a^2\big)$$

$$= -(1-v^2)\frac{(\lambda')^2}{2}\varphi_0(\varphi_0^2-a^2) + \frac{\lambda^2}{2}\varphi_0(\varphi_0^2-a^2) = \Big(-\frac{\lambda^2}{2}+\frac{\lambda^2}{2}\Big)\varphi_0(\varphi_0^2-a^2) = 0 \ . \tag{10.6.14}$$

Isto comprova que (10.6.11) é uma onda solitária, isto é, uma solução viajante para o campo φ que se translada com velocidade constante sem mudança de forma.

A densidade de energia da onda solitária (10.6.11) é

$$\mathcal{H} = \frac{1}{2}\left[\Big(\frac{\partial\varphi}{\partial t}\Big)^2 + \Big(\frac{\partial\varphi}{\partial x}\Big)^2\right] + \frac{\lambda^2}{8}(\varphi^2-a^2)^2 = \frac{v^2}{2}\varphi_0'(\xi;\lambda')^2 + \frac{1}{2}\varphi_0'(\xi;\lambda')^2 + \frac{\lambda^2}{8}(\varphi_0(\xi;\lambda')^2-a^2)^2$$

$$= \frac{1+v^2}{2}\varphi_0'(\xi;\lambda')^2 + \frac{\lambda^2}{8}\left(\varphi_0(\xi;\lambda')^2 - a^2\right)^2 = \frac{1+v^2}{2}\frac{(\lambda')^2}{4}\left(\varphi_0^2 - a^2\right)^2 + \frac{\lambda^2}{8}\left(\varphi_0^2 - a^2\right)^2$$

$$= \frac{\lambda^2}{8}\Big(1 + \frac{1+v^2}{1-v^2}\Big)\left(\varphi_0^2 - a^2\right)^2 = \frac{1}{4}\frac{\lambda^2}{1-v^2}\left[\varphi_0(\xi;\lambda')^2 - a^2\right]^2 \ . \qquad (10.6.15)$$

onde usamos (10.6.6) com λ' no lugar de λ. Assim,

$$\mathcal{H} = \frac{1}{4}\frac{\lambda^2 a^4}{1-v^2}\,\mathrm{sech}^4\left[\lambda a\gamma(x-vt)/2\right] \ , \qquad (10.6.16)$$

e, com a mudança de variáveis $y = \lambda a\gamma(x-vt)/2$, obtemos facilmente

$$E = \int_{-\infty}^{\infty}\mathcal{H}dx = \frac{\sqrt{1-v^2}}{1-v^2}\frac{\lambda a^3}{2}\int_{-\infty}^{\infty}\frac{dy}{\cosh^4 y} = \frac{2\lambda a^3/3}{\sqrt{1-v^2}} = \frac{E_0}{\sqrt{1-v^2}} \ , \qquad (10.6.17)$$

onde usamos (10.6.10). Este resultado mostra que a onda solitária se comporta como uma partícula relativística.

■ **Exercício 10.6.1.** Determine o momento linear transportado pela onda solitária (10.6.11) e interprete fisicamente o resultado obtido. ■

Depois de se chocarem, os "kinks" da teoria descrita pela lagrangiana (10.6.1) não preservam sua forma. Ondas solitárias que sempre se recompõem após colidirem entre si são conhecidas como *sólitons*. A chamada equação de "sine-Gordon" (Problema 10.8) dá origem aos exemplos mais simples e notáveis de sólitons em uma dimensão espacial.

10.7 Campos Vinculados

Se as Eqs.(10.4.1) não forem mutuamente independentes, elas não podem ser resolvidas para todas as velocidades $\dot{\varphi}_\alpha$ em termos dos momentos π^α e aparecem vínculos, isto é, relações de dependência entre os campos e seus momentos conjugados. Com adaptações até certo ponto óbvias, o formalismo desenvolvido na Seção 8.8 estende-se naturalmente às teorias de campos.

Campo de Schrödinger

Consideremos, à guisa de ilustração, a lagrangiana

$$\mathcal{L} = i\hbar\psi^*\dot{\psi} - \frac{\hbar^2}{2m}\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\psi^* \qquad (10.7.1)$$

que, conforme o Exercício 10.1.2, descreve o campo de Schrödinger livre.[10] Tomando ψ e ψ^* como campos independentes, os momentos conjugados correspondentes são

$$\pi(x) = \frac{\partial\mathcal{L}}{\partial\dot{\psi}(x)} = i\hbar\psi^*(x) \quad , \quad \pi_*(x) = \frac{\partial\mathcal{L}}{\partial\dot{\psi}^*(x)} = 0 \ . \qquad (10.7.2)$$

[10] Este exemplo é instrutivo porque exibe as principais características do campo de Dirac, que descreve partículas de spin $1/2$, sem as complicações algébricas inerentes a um campo com quatro componentes.

A hamiltoniana é dada por

$$\mathcal{H} = \pi\,\dot{\psi} + \pi_*\,\dot{\psi}^* - \mathcal{L} = \frac{\hbar^2}{2m}\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\psi^* \ . \tag{10.7.3}$$

O parêntese de Poisson é definido por

$$\{X,Y\} = \int d^3x \left(\frac{\delta X}{\delta\psi(x)}\frac{\delta Y}{\delta\pi(x)} - \frac{\delta X}{\delta\pi(x)}\frac{\delta Y}{\delta\psi(x)} + \frac{\delta X}{\delta\psi^*(x)}\frac{\delta Y}{\delta\pi_*(x)} - \frac{\delta X}{\delta\pi_*(x)}\frac{\delta Y}{\delta\psi^*(x)} \right) , \tag{10.7.4}$$

e os parênteses de Poisson fundamentais diferentes de zero são

$$\{\psi(\mathbf{x},t),\pi(\mathbf{y},t)\} = \delta(\mathbf{x}-\mathbf{y}) \quad , \quad \{\psi^*(\mathbf{x},t),\pi_*(\mathbf{y},t)\} = \delta(\mathbf{x}-\mathbf{y}) \ . \tag{10.7.5}$$

As velocidades não aparecem em (10.7.2), de modo que a teoria possui os vínculos primários

$$\phi_1(x) \equiv \pi(x) - i\hbar\psi^*(x) \approx 0 \quad , \quad \phi_2(x) \equiv \pi_*(x) \approx 0 \ . \tag{10.7.6}$$

Por analogia com a Eq.(8.8.6), na presença de vínculos as equações de movimento são geradas pela hamiltoniana total

$$H_T = H + \int d^3x \left[\lambda_1(x)\phi_1(x) + \lambda_2(x)\phi_2(x)\right] = \int d^3x \left\{ \mathcal{H} + \lambda_1(x)\phi_1(x) + \lambda_2(x)\phi_2(x) \right\}$$

$$= \int d^3x \left\{ \frac{\hbar^2}{2m}\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\psi^* + \lambda_1(x)\big[\pi(x) - i\hbar\psi^*(x)\big] + \lambda_2(x)\pi_*(x) \right\} \ . \tag{10.7.7}$$

De (10.7.5) deduz-se

$$\{\phi_1(\mathbf{x},t),\phi_2(\mathbf{y},t)\} = -i\hbar\,\{\psi^*(\mathbf{x},t),\pi_*(\mathbf{y},t)\} = -i\hbar\,\delta(\mathbf{x}-\mathbf{y}) \ , \tag{10.7.8}$$

de modo que a matriz formada pelos parênteses de Poisson dos vínculos tem elementos

$$\{\phi_r(\mathbf{x},t),\phi_s(\mathbf{y},t)\} = -i\hbar\,\epsilon_{rs}\,\delta(\mathbf{x}-\mathbf{y}) \ , \tag{10.7.9}$$

onde

$$\epsilon = (\epsilon_{rs}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \ . \tag{10.7.10}$$

A matriz (10.7.9) tem índices discretos e índices contínuos e, por analogia com (8.8.16), sua inversa, caso exista, tem elementos $C_{rs}(\mathbf{x},\mathbf{y})$ definidos por

$$\int d^3z \sum_{j=1}^{2} C_{rj}(\mathbf{x},\mathbf{z})\,\{\phi_j(\mathbf{z},t),\phi_s(\mathbf{y},t)\} = \delta_{rs}\,\delta(\mathbf{x}-\mathbf{y}) \ . \tag{10.7.11}$$

Esta definição justifica-se porque, no que se refere aos índices contínuos, os elementos da "matriz" identidade são $\delta(\mathbf{x}-\mathbf{y})$. Notando que $\boldsymbol{\epsilon}^2 = -\boldsymbol{I}$, verifica-se facilmente que (10.7.11) é satisfeita por

$$C_{rs}(\mathbf{x},\mathbf{y}) = \frac{1}{i\hbar}\,\epsilon_{rs}\,\delta(\mathbf{x}-\mathbf{y}) \ . \tag{10.7.12}$$

A existência da inversa de (10.7.9) estabelece que os vínculos são de segunda classe. Introduzindo os parênteses de Dirac

$$\{X,Y\}^* = \{X,Y\} - \sum_{rs}\int d^3x d^3y \{X,\phi_r(\mathbf{x},t)\}C_{rs}(\mathbf{x},\mathbf{y})\{\phi_s(\mathbf{y},t),Y\} \tag{10.7.13}$$

ou, explicitamente,

$$\{X,Y\}^* = \{X,Y\} - \frac{1}{i\hbar}\int d^3x\Big[\{X,\phi_1(\mathbf{x},t)\}\{\phi_2(\mathbf{x},t),Y\} - \{X,\phi_2(\mathbf{x},t)\}\{\phi_1(\mathbf{x},t),Y\}\Big] \ , \tag{10.7.14}$$

os vínculos (10.7.6) podem ser tomados como equações fortes e a hamiltoniana total (10.7.7) reduz-se à hamiltoniana canônica

$$H = \int d^3x\,\frac{\hbar^2}{2m}\nabla\psi\cdot\nabla\psi^* \ . \tag{10.7.15}$$

Os momentos canônicos são eliminados da teoria por intermédio das equações de vínculo (10.7.6) e os parênteses de Dirac fundamentais são

$$\{\psi(\mathbf{x},t),\psi(\mathbf{y},t)\}^* = \{\psi^*(\mathbf{x},t),\psi^*(\mathbf{y},t)\}^* = 0 \ , \tag{10.7.16a}$$

$$\{\psi(\mathbf{x},t),\psi^*(\mathbf{y},t)\}^* = \frac{1}{i\hbar}\,\delta(\mathbf{x}-\mathbf{y}) \ . \tag{10.7.16b}$$

■ **Exercício 10.7.1.** Usando (10.7.14), verifique as equações (10.7.16). ■

De acordo com as regras de quantização de sistemas com vínculos de segunda classe, devemos substituir o parêntese de Dirac por $i\hbar$ vezes o comutador (para campos bosônicos). Assim, o comutador fundamental dos operadores de campo $\hat{\Psi}$ e $\hat{\Psi}^\dagger$ escreve-se

$$[\hat{\Psi}(\mathbf{x},t),\hat{\Psi}^\dagger(\mathbf{y},t)] = \delta(\mathbf{x}-\mathbf{y}) \ , \tag{10.7.17}$$

onde $\hat{\Psi}^\dagger$ é o adjunto (conjugado hermitiano) de $\hat{\Psi}$. É o comutador (10.7.17) que, no

caso bosônico, serve de base à chamada segunda quantização da mecânica quântica não-relativística (Schiff 1968). A descrição de férmions requer a substituição do comutador pelo anticomutador em (10.7.17).

Campo Eletromagnético

O campo eletromagnético livre é outro sistema em cuja formulação hamiltoniana aparecem vínculos. Da Equação (10.2.13), com $\mu = 0$, deduz-se que os momentos canônicos são

$$\pi^\alpha = \frac{\partial \mathcal{L}}{\partial \dot{A}_\alpha} = F^{\alpha\, 0} \ , \tag{10.7.18}$$

ou, explicitamente,

$$\pi^0 = 0 \quad , \quad \pi^i = F^{i\, 0} = \partial^i A^0 - \dot{A}^i = E^i \ , \tag{10.7.19}$$

onde E^i, $i = 1, 2, 3$, são as componentes cartesianas do campo elétrico. Há, portanto, um vínculo primário

$$\phi_1 \equiv \pi^0 \approx 0 \ . \tag{10.7.20}$$

Um cálculo direto mostra que a hamiltoniana canônica toma a forma

$$H = \int d^3x \left(\frac{1}{2} \pi^i \pi^i + \frac{1}{4} F^{ij} F_{ij} + \pi^i \partial_i A_0 \right) \ . \tag{10.7.21}$$

Com o emprego dos parênteses de Poisson fundamentais

$$\{ A_\mu(\mathbf{x}, t), \pi^\nu(\mathbf{y}, t) \} = \delta_\mu^{\ \nu}\, \delta(\mathbf{x} - \mathbf{y}) \ , \tag{10.7.22}$$

a condição de consistência $\dot{\pi}^0 \approx 0$ dá lugar ao vínculo secundário

$$\phi_2 \equiv -\partial_i \pi^i \approx 0 \ . \tag{10.7.23}$$

Este vínculo não é outra coisa senão $\boldsymbol{\nabla} \cdot \mathbf{E} = 0$, daí ser conhecido como vínculo da lei de Gauss. Não há outros vínculos porque ϕ_2 tem parêntese de Poisson nulo tanto com H quanto com ϕ_1.

■ **Exercício 10.7.2.** Usando os parênteses de Poisson fundamentais, mostre que $\{ \phi_2(\mathbf{x}, t), H \} = 0$ e $\{ \phi_1(\mathbf{x}, t), \phi_2(\mathbf{y}, t) \} = 0$. Sugestão: $\{ \partial_i\, \pi^i(\mathbf{x}, t), X \} = \partial_i \{ \pi^i(\mathbf{x}, t), X \}$. ■

A hamiltoniana total é

$$H_T = \int d^3x \left(\frac{1}{2} \pi^i \pi^i + \frac{1}{4} F^{ij} F_{ij} - A_0\, \partial_i\, \pi^i + \lambda_1 \phi_1 \right) \ , \tag{10.7.24}$$

tendo sido realizada uma integração por partes com o abandono de um termo de superfície,

que não afeta as equações de movimento. Os vínculos ϕ_1 e ϕ_2 são de primeira classe e a dinâmica envolve funções arbitrárias, como A_0, refletindo a invariância da teoria sob transformações de calibre. A eliminação das funções arbitrárias requer a fixação do calibre. Deve-se notar que A_0 desempenha o papel de um multiplicador de Lagrange para a lei de Gauss, de modo que as variáveis canônicas genuínas são as componentes espaciais do potencial vetor e seus momentos conjugados. Outros aspectos e detalhes adicionais do tratamento hamiltoniano da eletrodinâmica de Maxwell são deixados para os tratados especializados (Dirac 1964; Henneaux & Teitelboim 1992).

PROBLEMAS

10.1. Prove que duas densidades lagrangianas que só diferem por uma divergência $\partial_\lambda \Phi^\lambda(\varphi_\alpha, x)$ geram as mesmas equações de movimento. Compare este resultado com o Teorema 2.3.1.

10.2. Diversos autores preferem a seguinte definição direta de derivada funcional:

$$\frac{\delta F}{\delta f(y)} = \lim_{\epsilon \to 0} \frac{F[f + \epsilon\, \delta_y] - F[f]}{\epsilon} \ ,$$

com $\delta_y(x) = \delta(x - y)$. Prove que esta definição é equivalente à da Seção 10.3.

10.3. Mostre que

$$\frac{\delta}{\delta f(x)} \int_{R^n} d^n y\, f(y)^2 = 2\, f(x) \ .$$

Usando este resultado e (10.3.4), prove que a equação diferencial funcional

$$\frac{\delta F}{\delta f(x)} + f(x) F = 0$$

admite como solução

$$F[f] = C \exp\Big\{-\frac{1}{2} \int_{R^n} d^n x\, f(x)^2\Big\} \ ,$$

onde C é uma constante arbitrária (Rosen 1969).

10.4. Dada a série de potências funcional

$$Z[f] = K_0 + \int d^n x_1\, K_1(x_1)\, f(x_1) + \frac{1}{2!} \int d^n x_1\, d^n x_2\, K_2(x_1, x_2)\, f(x_1)\, f(x_2)$$

$$+ \frac{1}{3!} \int d^n x_1\, d^n x_2\, d^n x_3\, K_3(x_1, x_2, x_3)\, f(x_1)\, f(x_2)\, f(x_3) + \cdots \ ,$$

onde $K_n(x_1, \ldots, x_n)$ é simétrica sob permutações de seus argumentos, prove que

$$K_n(x_1, \ldots, x_n) = \left. \frac{\delta^n Z}{\delta f(x_1) \cdots \delta f(x_n)} \right|_{f=0} \ .$$

Esta fórmula é de grande utilidade na teoria quântica de campos.

10.5. O tensor corrente de momento angular (orbital) $\mathcal{M}^{\alpha\beta\gamma}$ é definido por

$$\mathcal{M}^{\alpha\beta\gamma} = T^{\alpha\beta} x^\gamma - T^{\alpha\gamma} x^\beta \ ,$$

onde $T^{\alpha\beta}$ é o tensor de energia-momento. (i) O tensor momento angular $L^{\beta\gamma} = \int \mathcal{M}^{0\beta\gamma} d^3x$ será conservado se $\partial_\alpha \mathcal{M}^{\alpha\beta\gamma} = 0$. Prove que esta última equação é satisfeita se e somente se o tensor de energia-momento é simétrico, isto é, $T^{\alpha\beta} = T^{\beta\alpha}$. (ii) Outra razão física

para se exigir a simetria do tensor de energia-momento é que ele atua como fonte do campo gravitacional nas equações de Einstein, as quais exigem que as fontes da gravitação sejam representadas por tensores simétricos. Caso o tensor de energia-momento canônico (10.5.26) não seja simétrico, pode-se provar (Barut 1980; Soper 1976) que é sempre possível simetrizá-lo pela adição de uma quadridivergência $\partial_\lambda \psi^{\lambda\mu\nu}$ onde $\psi^{\lambda\mu\nu}$ é um tensor anti-simétrico nos dois primeiros índices λ, μ. Prove que se $\partial_\mu T^{\mu\nu} = 0$, também é verdade que $\partial_\mu \Theta^{\mu\nu} = 0$ com $\Theta^{\mu\nu} = T^{\mu\nu} + \partial_\lambda \psi^{\lambda\mu\nu}$. Sugestão: vide item (ii) do Problema 6.6. (iii) Desde que $\psi^{\lambda\mu\nu}$ se anule com suficiente rapidez no infinito, prove que as quantidades físicas conservadas correspondentes a $\Theta^{\mu\nu}$ e $T^{\mu\nu}$ têm o mesmo valor, o que justifica tomar $\Theta^{\mu\nu}$ como o tensor de energia-momento físico.

10.6. (i) Mostre que, de acordo com (10.5.26), o tensor de energia-momento canônico do campo eletromagnético livre é

$$T^{\mu\nu} = F^{\alpha\mu}\partial^\nu A_\alpha + \frac{1}{4} g^{\mu\nu} F^{\alpha\beta} F_{\alpha\beta} \ .$$

Note que este tensor não é simétrico. Outra objeção física a este tensor é que ele não é invariante sob transformações de calibre. (ii) Substituindo $\partial^\nu A_\alpha = \partial^\alpha A_\nu + F^\nu{}_\alpha$ na expressão de $T^{\mu\nu}$ e usando as equações da Maxwell para o campo eletromagnético livre, mostre que o tensor

$$\Theta^{\mu\nu} = F^{\alpha\mu} F^\nu{}_\alpha + \frac{1}{4} g^{\mu\nu} F^{\alpha\beta} F_{\alpha\beta}$$

é simétrico, invariante sob transformações de calibre e só difere de $T^{\mu\nu}$ por um termo da forma descrita no item (ii) do problema anterior. Exceto pelo fato de estar expresso em unidades de Heaviside-Lorentz, $\Theta^{\mu\nu}$ coincide com o tensor de energia-momento introduzido no Problema 6.18.

10.7. Considere um campo escalar em uma dimensão espacial descrito pela lagrangiana $\mathcal{L} = \frac{1}{2}\partial_\mu \partial^\mu \phi - V(\phi) \equiv \frac{1}{2}(\dot{\phi}^2 - \phi'^2) - V(\phi)$. (i) Se ϕ é uma solução estática da equação de Lagrange, prove que

$$\frac{d}{dx}\left[-\frac{1}{2}\phi'^2 + V(\phi)\right] = 0 \ .$$

(ii) Suponha $V(\phi) \geq 0$ e escreva $V(\phi)$ em termos de uma função $W(\phi)$ na forma $V(\phi) = \frac{1}{2}\left(\frac{dW}{d\phi}\right)^2$. Se ϕ é uma solução estática que para $x \to \pm\infty$ converge para alguma solução da equação algébrica $V(\phi) = 0$, mostre que ϕ satisfaz a equação diferencial de primeira ordem

$$\frac{d\phi}{dx} = \pm \frac{dW}{d\phi} \ .$$

(iii) Mostre que a energia $E = \int_{-\infty}^{\infty} dx[\frac{1}{2}\phi'^2 + V(\phi)]$ para uma solução estática do tipo acima pode ser escrita na forma

$$E = \frac{1}{2}\int_{-\infty}^{\infty} dx \left[\phi' \mp \frac{dW}{d\phi}\right]^2 \pm \Big[W(\phi(\infty)) - W(\phi(-\infty))\Big] = \pm\Big[W(\phi(\infty)) - W(\phi(-\infty))\Big] \ .$$

(iv) Aplique este resultado ao exemplo discutido na Seção 10.6 e obtenha a energia do "kink" (10.6.8) sem a necessidade de calcular explicitamente a integral (10.6.10).

10.8. Repita a análise da Seção 10.6 para a lagrangiana

$$\mathcal{L} = \frac{1}{2}\,\partial^\mu\phi\partial_\mu\phi - m^2\,(1-\cos\phi)$$

considerando a solução estática que conecta os estados fundamentais $\phi = 0$ e $\phi = 2\pi$. A equação de campo correspondente a esta lagrangiana é conhecida como equação de sine-Gordon.

10.9. Uma corda com extremos fixos em $x = 0$ e $x = \ell$ tem densidade linear de massa σ e tensão τ constantes. Se $\varphi_1(x,t)$ e $\varphi_2(x,t)$ são pequenos deslocamentos transversais da corda, ao longo de duas direções mutuamente perpendiculares, a densidade lagrangiana do sistema é

$$\mathcal{L} = \frac{\sigma}{2}\left[\left(\frac{\partial\varphi_1}{\partial t}\right)^2 + \left(\frac{\partial\varphi_2}{\partial t}\right)^2\right] - \frac{\tau}{2}\left[\left(\frac{\partial\varphi_1}{\partial x}\right)^2 + \left(\frac{\partial\varphi_2}{\partial x}\right)^2\right] .$$

(i) Obtenha as equações de Lagrange para $\varphi_1(x,t)$ e $\varphi_2(x,t)$. (ii) Levando em conta que $\varphi_1(x,t)$ e $\varphi_2(x,t)$ anulam-se nos extremos $x = 0$ e $x = \ell$, prove que a quantidade

$$Q = \int_0^\ell \sigma(\varphi_1\dot{\varphi}_2 - \dot{\varphi}_1\varphi_2)dx$$

é conservada, isto é, seu valor não depende do tempo. (iii) Que quantidade física Q representa? (iv) Para que valor de ω a onda estacionária $\varphi_1(x,t) = A\cos\omega t\,\mathrm{sen}\,(\pi x/\ell)$, $\varphi_2 = 0$ é solução das equações de Lagrange? Calcule a energia associada a esta onda.

10.10. O sistema formado por um campo escalar carregado em interação com o campo eletromagnético é descrito pela lagrangiana

$$\mathcal{L} = \left(\partial^\mu\phi + ieA^\mu\phi\right)\left(\partial_\mu\phi^* - ieA_\mu\phi^*\right) - \frac{1}{4}\,F^{\mu\nu}F_{\mu\nu} \ ,$$

onde e denota a carga da partícula descrita pelo campo ϕ. (i) Obtenha as equações de Lagrange. (ii) Prove que a lagrangiana é invariante sob a *transformação de calibre local* (também conhecida como *transformação de calibre de segunda espécie*)

$$\phi \to \tilde{\phi} = e^{-ie\chi}\phi \quad , \quad \phi^* \to \tilde{\phi}^* = e^{ie\chi}\phi^* \quad , \quad A_\mu \to \tilde{A}_\mu = A_\mu + \partial_\mu\chi \ ,$$

onde χ é uma função arbitrária. (iii) Mostre que a Lagrangiana pode ser escrita na forma

$$\mathcal{L} = \partial^\mu\phi\partial_\mu\phi^* - \frac{1}{4}\,F^{\mu\nu}F_{\mu\nu} - J^\mu A_\mu \ ,$$

onde J^μ é a corrente de Noether (10.5.33) associada á invariância da ação sob transformações de calibre globais, isto é, com χ constante.

10.11. As equações de Maxwell pressupõem que o fóton tem massa zero. Se a lei de Coulomb não for exatamente válida e o campo de uma carga puntiforme tiver um alcance finito, isto pode ser atribuído ao fato de o fóton possuir uma massa. Aparentemente, quem primeiro considerou as conseqüências de se atribuir uma massa ao fóton foi Proca. A lagrangiana de Proca é

$$\mathcal{L}_{\text{Proca}} = -\frac{1}{4}F_{\alpha\beta}F^{\alpha\beta} + \frac{\mu^2}{2}A_\alpha A^\alpha - \frac{1}{c}J_\alpha A^\alpha \ ,$$

onde o parâmetro μ tem dimensão de inverso de um comprimento. (i) Mostre que as equações de Lagrange para o campo de Proca A^α são

$$\partial_\beta F^{\beta\alpha} + \mu^2 A^\alpha = \frac{1}{c}J^\alpha \ ,$$

que não são invariantes sob transformações de calibre $A_\alpha \rightarrow A_\alpha + \partial_\alpha \Lambda$. (ii) Mostre que a conservação local da carga elétrica exige que o campo de Proca satisfaça a condição de Lorentz $\partial_\alpha A^\alpha = 0$. Prove que, como conseqüência, o campo de Proca obedece à seguinte equação:

$$\Box A^\alpha + \mu^2 A^\alpha = \frac{1}{c}J^\alpha \ .$$

No caso estático, em que há apenas uma carga puntiforme em repouso na origem, mostre que a solução esfericamente simétrica para o campo de Proca é $A^\alpha = (\phi, 0, 0, 0)$ com

$$\phi(r) = \frac{q}{4\pi}\frac{e^{-\mu r}}{r} \ .$$

Na teoria de Proca o potencial de uma carga puntiforme decai exponencialmente com a distância e μ^{-1} dá uma medida do alcance do potencial. (iii) Supondo soluções na forma de ondas planas $A^\alpha(\mathbf{r}, t) = A^\alpha_{(0)} \exp(i\mathbf{k}\cdot\mathbf{r} - i\omega t)$ no vácuo ($J^\alpha = 0$), mostre que

$$\omega^2 = c^2k^2 + \mu^2c^2 \ .$$

Em termos de $E = \hbar\omega$ e $\mathbf{p} = \hbar\mathbf{k}$, isto representa a relação relativística entre energia e momento linear de uma partícula de massa $\mu\hbar/c$.

10.12. Num espaço-tempo com apenas duas dimensões espaciais há uma variante da teoria eletromagnética de Maxwell, conhecida como teoria de Chern-Simons, em que os fótons têm massa mas a invariância de calibre é preservada (Deser, Jackiw & Templeton 1982). A lagrangiana de Chern-Simons é

$$\mathcal{L}_{\text{CS}} = -\frac{1}{4}F_{\alpha\beta}F^{\alpha\beta} + \frac{\mu}{4}\epsilon^{\alpha\beta\gamma}F_{\alpha\beta}A_\gamma \ ,$$

onde $F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha$, μ é o parâmetro de Chern-Simons e $\epsilon^{\alpha\beta\gamma}$ é o símbolo de Levi-Civita tridimensional com $\epsilon^{012} = 1$; os índices gregos assumem apenas os valores $0, 1, 2$. (i) Mostre que sob a transformação de calibre $A_\alpha \rightarrow A_\alpha + \partial_\alpha\Lambda$ a lagrangiana $\mathcal{L}_{\text{CS}}$ varia apenas pela divergência $\partial_\gamma[(\mu/4)\,\epsilon^{\alpha\beta\gamma}F_{\alpha\beta}\,\Lambda]$. (ii) Mostre que as equações de Lagrange para o campo de Chern-Simons A_α são

$$\partial_\beta F^{\beta\alpha} + \frac{\mu}{2}\epsilon^{\alpha\beta\gamma}F_{\beta\gamma} = 0 \ ,$$

as quais são invariantes sob transformações de calibre. Explique esta propriedade à luz do Problema 10.1. (iii) Mostre que

$$(F^{\alpha\beta}) = \begin{pmatrix} 0 & -E_x & -E_y \\ E_x & 0 & -B \\ E_y & B & 0 \end{pmatrix} \ .$$

Num mundo bidimensional o campo magnético é um escalar. Como explicar este fenômeno? (iv) Mostre que, em termos de $\mathbf{E}$ e B, as equações de Chern-Simons completas, incluindo as equações homogêneas $\partial^\alpha F^{\beta\gamma} + \partial^\beta F^{\gamma\alpha} + \partial^\gamma F^{\alpha\beta} = 0$, são

$$\nabla\cdot\mathbf{E} = \mu B \ , \quad \frac{1}{c}\frac{\partial E_x}{\partial t} + \mu E_y = \frac{\partial B}{\partial y} \ , \quad \frac{1}{c}\frac{\partial E_y}{\partial t} - \mu E_x = -\frac{\partial B}{\partial x} \ , \quad \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -\frac{1}{c}\frac{\partial B}{\partial t} \ .$$

(v) Prove que $\mathbf{E}$ e B satisfazem a equação de Klein-Gordon $(\Box+\mu^2)\phi = 0$, de modo que μ faz o papel de massa do fóton, a partícula descrita pelo campo eletromagnético. (vi) Generalize a teoria de Chern-Simons para o caso em que há cargas e correntes, descritas pela densidade de corrente J^α. Prove que, assim como as equações de Maxwell, as equações da teoria de Chern-Simons implicam automaticamente a conservação local da carga elétrica, isto é, $\partial_\alpha J^\alpha = 0$.

10.13. Considere uma teoria de campos com tensor de energia-momento simétrico, de modo que o quadrimomento P^μ e o tensor momento angular $L^{\mu\nu}$ são conservados. Mostre que as componentes L^{0i} do tensor momento angular podem ser escritas na forma

$$L^{0i} = P^i x^0 - P^0 x^i_c \ ,$$

onde P^0 é a energia total e x^i_c são as componentes do vetor posição do *centro de energia*:

$$x^i_c = \frac{\int d^3x T^{00} x^i}{\int d^3x T^{00}} \ .$$

Conclua que, para um sistema relativístico, a lei de conservação $dL^{0i}/dx^0 = 0$ equivale a

$$\frac{dx^i_c}{dx^0} = \frac{P^i}{P^0} \ .$$

Interprete este resultado comparando-o com a Eq.(1.1.9) da mecânica newtoniana.

Apêndice A

Notação Indicial

Este apêndice discute as regras básicas para a manipulação correta de somas envolvendo quantidades indexadas.

A soma

$$s = a_1 + a_2 + \ldots + a_N \tag{A.1}$$

pode ser escrita numa forma compacta com o emprego do sinal de somatório:

$$s = \sum_{i=1}^{N} a_i \ . \tag{A.2}$$

É evidente que as equações a seguir têm o mesmo significado que a Eq.(A.2):

$$s = \sum_{j=1}^{N} a_j \ , \tag{A.3}$$

$$s = \sum_{k=1}^{N} a_k \ , \tag{A.4}$$

e assim por diante. O índice i na Eq.(A.2), j na Eq.(A.3) ou k na Eq.(A.4) é chamado de índice "mudo", significando que a soma é independente da letra usada para indicá-la.

Uma soma dupla da forma

$$\sigma = \sum_{i,j=1}^{N} b_{ij} \equiv \sum_{i=1}^{N} \sum_{j=1}^{N} b_{ij} \ , \tag{A.5}$$

significa explicitamente

$$\sigma = b_{11} + b_{12} + \ldots + b_{1N} + b_{21} + b_{22} + \ldots + b_{2N} + \ldots + b_{N1} + b_{N2} + \ldots + b_{NN} \ . \tag{A.6}$$

É claro que

$$\sigma = \sum_{i,j=1}^{N} b_{ji} = \sum_{j,i=1}^{N} b_{ij} = \sum_{j,i=1}^{N} b_{ji} \ , \tag{A.7}$$

e, ainda,

$$\sigma = \sum_{k,l=1}^{N} b_{kl} = \sum_{m,n=1}^{N} b_{mn} \ , \tag{A.8}$$

et cetera, toda estas somas representando exatamente a mesma quantidade expressa pela Eq.(A.6). Combinando as Eqs.(A.5) e (A.7) obtemos um resultado útil:

$$\sum_{i,j=1}^{N} b_{ij} = \frac{1}{2} \sum_{i,j=1}^{N} (b_{ij} + b_{ji}) \ . \tag{A.9}$$

Somas triplas, quádruplas, etc, são definidas como generalizações imediatas da Eq.(A.5).

O sistema de equações

$$\begin{aligned} y_1 &= a_{11}\, x_1 + a_{12}\, x_2 + \ldots + a_{1N}\, x_N \ , \\ y_2 &= a_{21}\, x_1 + a_{22}\, x_2 + \ldots + a_{2N}\, x_N \ , \\ &\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ y_M &= a_{M1}\, x_1 + a_{M2}\, x_2 + \ldots + a_{MN}\, x_N \ , \end{aligned} \tag{A.10}$$

pode ser reescrito na forma abreviada

$$y_i = \sum_{j=1}^{N} a_{ij}\, x_j \quad , \quad i = 1, \ldots, M \ . \tag{A.11}$$

Um índice que aparece apenas uma vez em cada termo da soma e que não indica nenhum somatório, como i na Eq.(A.11), é chamado de índice "livre". Note que

$$y_j = \sum_{l=1}^{N} a_{jl}\, x_l \quad , \quad j = 1, \ldots, M \tag{A.12}$$

ou

$$y_k = \sum_{m=1}^{N} a_{km}\, x_m \quad , \quad k = 1, \ldots, M \tag{A.13}$$

representam exatamente a mesma coisa que a Eq.(A.11). É claro que o índice livre que aparece em cada termo de uma equação deve ser o mesmo. Assim, as equações

$$a_i + b_i = c_i \ , \tag{A.14}$$

$$a_i + \sum_{j=1}^{N} b_{ij}\, c_j = d_i \tag{A.15}$$

fazem sentido, ao passo que

$$a_i = b_j \tag{A.16}$$

de modo geral é uma equação desprovida de significado.

Delta de Kronecker

O símbolo delta de Kronecker, denotado por δ_{ij}, é definido por

$$\delta_{ij} = \begin{cases} 1 & \text{se} \quad i = j \\ 0 & \text{se} \quad i \neq j \end{cases} \quad . \tag{A.17}$$

A principal propriedade do delta de Kronecker é a seguinte:

$$\sum_{l=1}^{N} \delta_{kl}\, a_l = a_k \quad , \quad k = 1, \ldots, N \ . \tag{A.18}$$

Esta propriedade é de demonstração imediata, pois

$$\sum_{l=1}^{N} \delta_{kl}\, a_l = \delta_{k1} a_1 + \ldots + \delta_{k\,k-1} a_{k-1} + \delta_{kk} a_k + \delta_{k\,k+1} a_{k+1} + \ldots + \delta_{kN} a_N \ . \tag{A.19}$$

Nesta última equação todos os termos são nulos exceto o termo $\delta_{kk} a_k = a_k$, ficando estabelecida a Eq.(A.18). Naturalmente, pelo mesmo argumento,

$$\sum_{j=1}^{N} \delta_{ij}\, T_{jk} = T_{ik} \ , \tag{A.20}$$

a mesma propriedade valendo se houver um número qualquer de índices livres.

Símbolo de Levi-Civita

O *símbolo de Levi-Civita* ϵ_{ijk} é completamente definido por $\epsilon_{123} = 1$ e pela propriedade de ser anti-simétrico sob qualquer troca de índices adjacentes. Em outras palavras, $\epsilon_{ijk} = \pm 1$ conforme (i, j, k) seja uma permutação par ou ímpar de $(1, 2, 3)$ e $\epsilon_{ijk} = 0$ caso haja repetição de índices. Usando notação indicial para as componentes cartesianas de vetores, como em $\boldsymbol{a} = (a_1, a_2, a_3)$, o produto vetorial $\mathbf{c} = \boldsymbol{a} \times \mathbf{b}$ pode ser expresso na forma

$$c_i = \sum_{j,k=1}^{3} \epsilon_{ijk}\, a_j\, b_k \ . \tag{A.21}$$

Uma identidade útil é

$$\sum_{k=1}^{3} \epsilon_{ijk}\, \epsilon_{klm} = \delta_{il}\, \delta_{jm} - \delta_{im}\, \delta_{jl} \ . \tag{A.22}$$

A demonstração deste último resultado consiste em observações simples. O primeiro membro só é diferente de zero se $i = l$ e $j = m$, ou se $i = m$ e $j = l$, logo pode ser expresso como uma combinação linear dos produtos de deltas de Kronecker que aparecem no segundo membro da identidade. Escolhas de valores particulares para os índices determinam os coeficientes.

Substituição e Fatoração

Há certos cuidados que precisam ser tomados durante manipulações algébricas envolvendo a notação indicial. Consideremos, inicialmente, o processo de *substituição*. Se

$$a_i = \sum_{k=1}^{N} U_{ik}\, b_k \tag{A.23}$$

e

$$b_i = \sum_{k=1}^{N} V_{ik}\, c_k \ , \tag{A.24}$$

para substituir em (A.23) os bs dados por (A.24) primeiro mudamos o índice livre em (A.24) de i para k e o índice mudo k para outra letra ainda não utilizada, digamos l, de modo que

$$b_k = \sum_{l=1}^{N} V_{kl}\, c_l \ . \tag{A.25}$$

Agora, a substituição de (A.25) em (A.23) fornece

$$a_i = \sum_{k,l=1}^{N} U_{ik}\, V_{kl}\, c_l \ . \tag{A.26}$$

Se a transição de (A.24) para (A.25) com a mudança do índice de soma não tivesse sido realizada, a troca pura e simples de i por k em (A.24) seguida de sua introdução direta em (A.23) resultaria em

$$a_i \stackrel{?}{=} \sum_{k=1}^{N} \sum_{k=1}^{N} U_{ik}\, V_{kk}\, c_k \ , \tag{A.27}$$

que é um resultado flagrantemente incorreto porque exclui termos não-diagonais, tais como V_{12}, V_{23}, que teriam que aparecer. Para a substituição redundar numa equação correta é *indispensável* que os índices de soma em (A.23) e (A.24) sejam *distintos*, uma vez que os respectivos somatórios são mutuamente independentes. A mesma precaução deve ser tomada na multiplicação de somatórios. Por exemplo, se

$$p = \sum_{k=1}^{N} a_k \tag{A.28}$$

e

$$q = \sum_{k=1}^{N} b_k \ , \tag{A.29}$$

e expressão correta para o produto pq é

$$pq = \sum_{k,l=1}^{N} a_k \, b_l \ . \tag{A.30}$$

Insistimos que

$$pq \neq \sum_{k=1}^{N} \sum_{k=1}^{N} a_k \, b_k = N \left(a_1 b_1 + a_2 b_2 + \ldots + a_N b_N \right) \ , \tag{A.31}$$

ou seja, o produto de duas somas só pode ser expresso corretamente como soma dupla se forem usados índices mudos distintos para representar os somatórios independentes a serem multiplicados, como na Eq.(A.30).

Finalmente, no que concerne ao processo de *fatoração*, o procedimento correto é indicado por meio de um exemplo. Considere a equação

$$\sum_{j=1}^{N} T_{ij} \, n_j - \lambda n_i = 0 \ , \tag{A.32}$$

na qual queremos fatorar n_j. Com a ajuda do delta de Kronecker escrevemos

$$n_i = \sum_{j=1}^{N} \delta_{ij} \, n_j \ , \tag{A.33}$$

que, inserida em (A.32), fornece

$$\sum_{j=1}^{N} T_{ij} \, n_j - \lambda \sum_{j=1}^{N} \delta_{ij} \, n_j = 0 \ , \tag{A.34}$$

ou

$$\sum_{j=1}^{N} \left(T_{ij} - \lambda \delta_{ij} \right) n_j = 0 \ , \tag{A.35}$$

conforme se pretendia.

Apêndice B

Funções Homogêneas e Teorema de Euler

Uma função $F : \boldsymbol{R}^m \to \boldsymbol{R}$ é dita *homogênea de grau* p se existe um número real p tal que

$$F(\lambda x_1, \ldots, \lambda x_m) = \lambda^p F(x_1, \ldots, x_m) \quad , \quad \lambda > 0 \ , \tag{B.1}$$

onde o número real positivo λ é arbitrário.

■ **Exemplo B.1.** A função $F(x,y) = (x^4 + 2xy^3 - 5y^4)\,\mathrm{sen}(x/y)$ é homogênea de grau 4, pois $F(\lambda x, \lambda y) = (\lambda^4 x^4 + 2\lambda x \lambda^3 y^3 - 5\lambda^4 y^4)\,\mathrm{sen}(\lambda x/\lambda y) = \lambda^4 F(x,y)$. Por sua vez, com A constante, a função $F(x,y,z) = A/(x^2+y^2+z^2)^{1/2}$ satisfaz $F(\lambda x, \lambda y, \lambda z) = A/(\lambda^2 x^2 + \lambda^2 y^2 + \lambda^2 z^2)^{1/2} = \lambda^{-1} F(x,y,z)$, logo é homogênea de grau -1. ■

Teorema de Euler. Se $F : \boldsymbol{R}^m \to \boldsymbol{R}$ é uma função diferenciável e homogênea de grau p, então

$$\sum_{k=1}^{m} x_k \frac{\partial F}{\partial x_k} = p\,F \ . \tag{B.2}$$

Inversamente, se F satisfaz (B.2) então F é homogênea de grau p.

Demonstração. Suponha F homogênea de grau p e defina $u_1 = \lambda x_1, \ldots, u_m = \lambda x_m$, de modo que

$$F(u_1, \ldots, u_m) = \lambda^p F(x_1, \ldots, x_m) \ . \tag{B.3}$$

Diferenciando esta última equação em relação a λ, com $x_1, \ldots, x_m$ fixos, resulta

$$\sum_{k=1}^{m} \frac{\partial F}{\partial u_k} \frac{\partial u_k}{\partial \lambda} = p\,\lambda^{p-1} F \quad \Longrightarrow \quad \sum_{k=1}^{m} x_k \frac{\partial F}{\partial u_k} = p\,\lambda^{p-1} F \ . \tag{B.4}$$

Fazendo $\lambda = 1$ em (B.4) deduz-se (B.2), pois agora $u_1 = x_1, \ldots, u_m = x_m$. Inversamente, suponha que F obedeça à equação diferencial (B.2) e considere a função

$$g(\lambda) = \lambda^{-p} F(\lambda x_1, \ldots, \lambda x_m) \tag{B.5}$$

com $x_1, \ldots, x_m$ fixos. Então, novamente com $u_1 = \lambda x_1, \ldots, u_m = \lambda x_m$, temos

$$\frac{dg}{d\lambda} = -p\,\lambda^{-p-1} F(u_1, \ldots, u_m) + \lambda^{-p} \sum_{k=1}^{m} \frac{\partial F}{\partial u_k} x_k = \lambda^{-p-1} \left[-p\, F(u_1, \ldots u_m) + \sum_{k=1}^{m} \frac{\partial F}{\partial u_k} u_k \right] = 0 \tag{B.6}$$

em virtude de (B.2). Portanto, $g(\lambda) =$ constante. Mas de (B.5) infere-se que $g(1) = F(x_1, \ldots, x_m)$, de modo que $g(\lambda) = g(1) = F(x_1, \ldots, x_m)$, o que prova (B.1) e estabelece que F é uma função homogênea de grau p. □

■ **Exemplo B.2.** Pelo teorema de Euler, a solução geral da equação diferencial de primeira ordem para $f(x, y)$

$$x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} = p\, f$$

pode ser escrita na forma

$$f(x, y) = x^p\, g\left(\frac{x}{y}\right) ,$$

onde $g : \boldsymbol{R} \rightarrow \boldsymbol{R}$ é uma função diferenciável arbitrária. ■

Apêndice C

Espaços Vetoriais e Operadores Lineares

Neste apêndice expomos em linhas gerais a teoria das transformações lineares em espaços vetoriais de dimensão finita. Supomos que o leitor esteja familiarizado com a definição de espaço vetorial e com as noções de dependência linear, base, matriz, traço e determinante (Halmos 1974; Hoffmann & Kunze 1971). Desigaremos por $\mathcal{F}$ o corpo dos números reais ou complexos, que chamaremos genericamente de *escalares*. Denotaremos os escalares por letras latinas, tais como a, b, c, e os vetores por letras gregas, como α, β, γ.

C.1 Operadores Lineares

Definição C.1. Seja $\mathcal{V}$ um espaço vetorial sobre o corpo $\mathcal{F}$. Um *operador linear* (ou uma *transformação linear*) A em $\mathcal{V}$ é uma regra que faz corresponder a cada vetor α de $\mathcal{V}$ um vetor $A\alpha$ de $\mathcal{V}$ de tal maneira que

$$A(c_1\alpha + c_2\beta) = c_1 A\alpha + c_2 A\beta \qquad (C.1)$$

para todos os α, β em $\mathcal{V}$ e todos os escalares c_1, c_2 em $\mathcal{F}$.

Para um operador linear A é sempre verdade que $A0 = 0$. O *operador identidade* I, definido por $I\alpha = \alpha$, é uma transformação linear em $\mathcal{V}$, o mesmo ocorrendo com o *operador nulo* 0, definido por $0\alpha = 0$.

Dados dois operadores lineares A e B, sua *soma* $S = A + B$ é o operador linear definido por $S\alpha = A\alpha + B\alpha$, ao passo que seu *produto* $P = AB$ é definido por $P\alpha = A(B\alpha)$. Em geral, o produto de operadores lineares não é comutativo. As propriedades algébricas da soma e produto de números reais permanecem válidas para operadores, com a exceção notável da comutatividade.

Definição C.2. Um operador linear A é dito *inversível* se possui as duas seguintes propriedades:

(a) A cada vetor β corresponde pelo menos um vetor α tal que $A\alpha = \beta$;

(b) Se $\alpha_1 \neq \alpha_2$ então $A\alpha_1 \neq A\alpha_2$.

Estas propriedades caracterizam A como uma aplicação injetora e sobrejetora, isto é, como uma bijeção de $\mathcal{V}$ em $\mathcal{V}$.

Se A é inversível pode-se definir um operador linear A^{-1}, chamado de *inverso* de A, conforme passamos a descrever. Se β é qualquer vetor, podemos, por (a), encontrar um vetor α tal que $A\alpha = \beta$. Além disso, por (b), α é univocamente determinado. Por definição, $A^{-1}\beta = \alpha$. Para provar que A^{-1} é linear, calculemos $A^{-1}(c_1\beta_1 + c_2\beta_2)$ com $A^{-1}\beta_1 = \alpha_1$ e $A^{-1}\beta_2 = \alpha_2$. Da linearidade de A segue-se que $A(c_1\alpha_1 + c_2\alpha_2) = c_1\beta_1 + c_2\beta_2$, de modo que $A^{-1}(c_1\beta_1 + c_2\beta_2) = c_1\alpha_1 + c_2\alpha_2 = c_1A^{-1}\beta_1 + c_2A^{-1}\beta_2$ e A^{-1} é linear. Da definição de inverso decorre imediatamente que

$$AA^{-1} = A^{-1}A = I \ . \tag{C.2}$$

Teorema C.1. Um operador linear A num espaço vetorial de dimensão finita $\mathcal{V}$ é inversível se e somente se $A\alpha = 0$ implica $\alpha = 0$.

Demonstração. Se A é inversível, combinando $A0 = 0$ com a propriedade (b) conclui-se que a única solução de $A\alpha = 0$ é $\alpha = 0$. Suponha, agora, que $A\alpha = 0$ implique $\alpha = 0$. Neste caso, $\alpha_1 \neq \alpha_2$ ou, equivalentemente, $\alpha_1 - \alpha_2 \neq 0$ implica $A(\alpha_1 - \alpha_2) \neq 0$, isto é, $A\alpha_1 \neq A\alpha_2$, o que prova (b). Para provar (a), seja $\{\alpha_1, \ldots, \alpha_n\}$ uma base de $\mathcal{V}$. Se $\sum_i c_iA\alpha_i = 0$ então $A(\sum_i c_i\alpha_i) = 0$ donde, conforme nossa hipótese, $\sum_i c_i\alpha_i = 0$ e a independência linear dos α_i conduz a $c_1 = \cdots = c_n = 0$. Portanto, $\{A\alpha_1, \ldots, A\alpha_n\}$ é um conjunto de vetores linearmente independentes, constituindo, também, uma base de $\mathcal{V}$, e qualquer vetor β pode ser escrito na forma $\beta = \sum_i c_iA\alpha_i = A(\sum_i c_i\alpha_i) \equiv A\alpha$, o que completa a demonstração. □

Verifica-se facilmente que: (i) se A e B são inversíveis o produto AB também é inversível e $(AB)^{-1} = B^{-1}A^{-1}$; (ii) se A é inversível e $c \neq 0$, então cA é inversível e $(cA)^{-1} = \frac{1}{c}A^{-1}$; (iii) se A é inversível, então A^{-1} também é inversível e $(A^{-1})^{-1} = A$.

C.2 Representação Matricial de Operadores

Definição C.3. Seja $\mathcal{V}$ um espaço vetorial de dimensão n, seja $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ uma base de $\mathcal{V}$ e seja A um operador linear em $\mathcal{V}$. Já que qualquer vetor é uma combinação linear dos α_i, temos

$$A\alpha_j = \sum_{i=1}^{n} a_{ij}\alpha_i \quad , \quad j = 1, \ldots, n \ . \tag{C.3}$$

O conjunto (a_{ij}) de n^2 escalares é a matriz de A na base (ou sistema de coordenadas) $\mathcal{B}$, que denotaremos por $[A]$. Se for necessário especificar a base sob consideração, escreveremos $[A]_{\mathcal{B}}$.

Sejam $\xi = \sum_i x_i\alpha_i$ e $\eta = \sum_i y_i\alpha_i$ vetores tais que $\eta = A\xi$. Sejam $[\xi]$ e $[\eta]$ as matrizes-coluna

$$[\xi] = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad , \quad [\eta] = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad . \tag{C.4}$$

Então

$$\eta = A\xi = \sum_{j=1}^{n} x_j A\alpha_j = \sum_{j=1}^{n}\sum_{i=1}^{n} x_j a_{ij}\alpha_i \quad , \tag{C.5}$$

donde

$$y_i = \sum_{j=1}^{n} a_{ij} x_j \quad , \tag{C.6}$$

ou, em linguagem matricial,

$$[\eta] = [A]\,[\xi] \quad . \tag{C.7}$$

É para garantir a validade desta equação que se faz a escolha de índices aparentemente perversa na Eq.(C.3). As operações algébricas entre operadores lineares transferem-se inalteradas para as matrizes correspondentes. Por exemplo, fica como um exercício para o leitor a comprovação de que se $C = AB$ então $[C] = [A]\,[B]$. Em particular, se S é um operador inversível, de $SS^{-1} = S^{-1}S = I$ segue-se que $[S]\,[S^{-1}] = [S^{-1}]\,[S] = \boldsymbol{I}$, onde $\boldsymbol{I}$ é a matriz identidade, e, conseqüentemente, $[S^{-1}] = [S]^{-1}$. Um operador linear é inversível se e somente se sua matriz associada é regular ou não-singular, isto é, tem determinante diferente de zero em alguma base (e, portanto, em todas as bases). Pode-se demonstrar (Hoffmann & Kunze 1971) que os elementos da inversa de uma matriz não-singular M são dados por

$$(M^{-1})_{ij} = (-1)^{i+j}\,\frac{\det M(j|i)}{\det M} \quad , \tag{C.8}$$

onde $M(i|j)$ é a matriz que se obtém apagando a i-ésima linha e a j-ésima coluna de M. No segundo membro de (C.8) deve-se atentar para a ordem invertida dos índices.

De que forma a matriz associada a um operador linear A é afetada por uma mudança de base? Para responder a esta pergunta, sejam $\mathcal{X} = \{\xi_1, \ldots, \xi_n\}$ e $\mathcal{Y} = \{\eta_1, \ldots, \eta_n\}$ duas bases de $\mathcal{V}$ e considere o operador linear de mudança de base S definido por $S\xi_i = \eta_i$, ou, mais explicitamente,

$$S\,(\sum_i x_i\,\xi_i) = \sum_i x_i\,\eta_i \quad . \tag{C.9}$$

O operador S é claramente inversível, pois $S(\sum_i x_i\,\xi_i) = 0$ equivale a $\sum_i x_i\,\eta_i = 0$, que implica $x_1 = \ldots = x_n = 0$. Sejam $[A]_{\mathcal{X}} = (a_{ij})$, $[A]_{\mathcal{Y}} = (b_{ij})$ e $[S]_{\mathcal{X}} = (s_{ij})$. Podemos escrever

$$A\xi_j = \sum_i a_{ij}\xi_i \tag{C.10}$$

e

$$A\eta_j = \sum_i b_{ij}\eta_i \ . \tag{C.11}$$

Portanto,

$$A\eta_j = AS\xi_j = A(\sum_k s_{kj}\xi_k) = \sum_k s_{kj}A\xi_k = \sum_k s_{kj}(\sum_i a_{ik}\xi_i) = \sum_i(\sum_k a_{ik}s_{kj})\xi_i \ , \tag{C.12}$$

ao mesmo tempo que

$$A\eta_j = \sum_k b_{kj}\eta_k = \sum_k b_{kj}S\xi_k = \sum_k b_{kj}(\sum_i s_{ik}\xi_i) = \sum_i(\sum_k s_{ik}b_{kj})\xi_i \ . \tag{C.13}$$

Comparando (C.12) com (C.13) vemos que

$$\sum_k a_{ik}s_{kj} = \sum_k s_{ik}b_{kj} \ , \tag{C.14}$$

isto é,

$$[A]_{\mathcal{X}}[S]_{\mathcal{X}} = [S]_{\mathcal{X}}[A]_{\mathcal{Y}} \ . \tag{C.15}$$

A forma mais usual de escrever esta equação é

$$[A]_{\mathcal{Y}} = [S]_{\mathcal{X}}^{-1}[A]_{\mathcal{X}}[S]_{\mathcal{X}} \ . \tag{C.16}$$

Duas matrizes M e N são ditas *similares* se estão relacionadas por uma *transformação de similaridade*, isto é, se existe uma matriz inversível P tal que $M = P^{-1}NP$. A Eq.(C.16) mostra que, sob uma mudança de base, a matriz associada a um operador linear sofre uma transformação de similaridade. Como $\mathrm{tr}\,(MN) = \mathrm{tr}\,(NM)$ e $\det(MN) = (\det M)(\det N)$, verifica-se prontamente que o traço e o determinante de uma matriz são invariantes sob uma transformação de similaridade. Assim, podemos falar em *traço de um operador* e *determinante de um operador* como quantidades independentes de qualquer escolha particular de base.

C.3 Autovalores e Autovetores

Definição C.4. Seja A um operador linear num espaço vetorial $\mathcal{V}$ sobre o corpo $\mathcal{F}$. Um *autovalor* de A é um escalar c tal que existe um vetor não-nulo α em $\mathcal{V}$ que satisfaz $A\alpha = c\alpha$. Dizemos que α é um *autovetor* de A associado ao autovalor c. O *espectro* de um operador linear A é o conjunto dos seus autovalores.

Teorema C.2. Todo operador linear sobre um espaço vetorial complexo de dimensão finita possui um autovetor.

Demonstração. Um operador linear A num espaço vetorial complexo $\mathcal{V}$ de dimensão n possui um autovetor se existe $\alpha \neq 0$ tal que $(A - cI)\alpha = 0$ para algum número complexo c. Logo, para que exista um autovetor é necessário e suficiente que o operador linear $A - cI$ não seja inversível, isto é, numa base qualquer de $\mathcal{V}$ devemos ter $\det([A] - c\boldsymbol{I}) = 0$ para algum número complexo c. Mas $\det([A] - c\boldsymbol{I})$ é um polinômio de grau n em c, e o teorema fundamental da álgebra assegura que todo polinômio de grau positivo com coeficientes complexos possui uma raiz complexa. Logo, existe um número complexo c tal que $\det([A] - c\boldsymbol{I}) = 0$ e a demonstração está completa. □

C.4 Produto Interno e Bases Ortonormais

Definição C.5. Seja $\mathcal{F}$ o corpo dos números reais ou o corpo dos números complexos e seja $\mathcal{V}$ um espaço vetorial sobre $\mathcal{F}$. Um *produto interno* ou *produto escalar* em $\mathcal{V}$ é uma regra que a cada par ordenado de vetores α, β de $\mathcal{V}$ associa o escalar (α, β) de tal maneira que:

(a) $(\alpha + \beta, \gamma) = (\alpha, \gamma) + (\beta, \gamma)$;
(b) $(\alpha, c\beta) = c(\alpha, \beta)$ para qualquer $c \in \mathcal{F}$;
(c) $(\alpha, \beta) = (\beta, \alpha)^*$, onde o asterisco indica conjugação complexa;
(d) $(\alpha, \alpha) > 0$ se $\alpha \neq 0$.

As propriedades (a), (b) e (c) têm as seguintes implicações imediatas:

(e) $(c\alpha, \beta) = c^*(\alpha, \beta)$; (f) $(\alpha, \beta + \gamma) = (\alpha, \beta) + (\alpha, \gamma)$.

Quando $\mathcal{F}$ é o corpo dos números reais, os complexos conjugados em (c) e (e) são supérfluos. Outra conseqüência direta da definição de produto interno é $(\alpha, 0) = 0$. Combinado com (d), este último resultado estabelece que $(\alpha, \alpha) = 0$ se e somente se $\alpha = 0$. Vale a pena mencionar que os matemáticos preferem substituir a propriedade (b) por (b') $(c\alpha, \beta) = c(\alpha, \beta)$.

Num espaço vetorial dotado de produto interno o número real não-negativo $||\alpha||$ definido por

$$||\alpha|| = \sqrt{(\alpha, \alpha)} \tag{C.17}$$

é chamado de *norma* ou *módulo* do vetor α. As seguintes propriedades da norma são de verificação imediata: (i) $||c\alpha|| = |c|\,||\alpha||$; (ii) $||\alpha|| \geq 0$ e $||\alpha|| = 0$ se e somente se $\alpha = 0$.

Sejam $\alpha = (x_1, \ldots, x_n)$ e $\beta = (y_1, \ldots, y_n)$ vetores de $\mathcal{F}^n$. O *produto interno canônico* em $\mathcal{F}^n$ é definido por

$$(\alpha, \beta) = x_1^* y_1 + x_2^* y_2 + \cdots + x_n^* y_n \ . \tag{C.18}$$

A *norma canônica* em $\mathcal{F}^n$ é dada por

$$||\alpha|| = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2} \ . \tag{C.19}$$

Teorema C.3. Num espaço vetorial com produto interno valem a *desigualdade de Schwarz*

$$|(\alpha, \beta)| \leq ||\alpha||\,||\beta|| \tag{C.20}$$

e a *desigualdade triangular*

$$||\alpha+\beta|| \leq ||\alpha|| + ||\beta|| \ . \tag{C.21}$$

Demonstração. Se $\alpha = 0$ a desigualdade (C.20) é obviamente verdadeira. Se $\alpha \neq 0$, considere $\gamma = \beta - c\alpha$. Usando as propriedades do produto interno podemos escrever

$$0 \leq (\gamma,\gamma) = ||\beta||^2 - c^*(\alpha,\beta) - c(\beta,\alpha) + |c|^2||\alpha||^2 \ . \tag{C.22}$$

A escolha $c = (\alpha,\beta)/||\alpha||^2$ conduz a

$$||\beta||^2 - \frac{|(\alpha,\beta)|^2}{||\alpha||^2} \geq 0 \ , \tag{C.23}$$

da qual decorre imediatamente a desigualdade de Schwarz. Usando agora (C.20), juntamente com a desigualdade evidente $\mathrm{Re}\, z \leq |z|$, temos

$$\begin{aligned} ||\alpha+\beta||^2 &= ||\alpha||^2 + (\alpha,\beta) + (\beta,\alpha) + ||\beta||^2 \\ &= ||\alpha||^2 + 2\mathrm{Re}\,(\alpha,\beta) + ||\beta||^2 \leq ||\alpha||^2 + 2||\alpha||\,||\beta|| + ||\beta||^2 = (||\alpha|| + ||\beta||)^2 \ , \end{aligned} \tag{C.24}$$

donde, por extração da raiz quadrada positiva, resulta a desigualdade triangular. □

Definição C.6. Num espaço vetorial $\mathcal{V}$ com produto interno, dois vetores α e β são ditos *ortogonais* se $(\alpha,\beta) = 0$, e escrevemos $\alpha \perp \beta$. Dois subconjuntos $\mathcal{R}$ e $\mathcal{S}$ de $\mathcal{V}$ são ditos ortogonais, e escrevemos $\mathcal{R} \perp \mathcal{S}$, se cada vetor de $\mathcal{R}$ é ortogonal a cada vetor de $\mathcal{S}$. Um conjunto de vetores tal que quaisquer dois vetores são ortogonais é dito um *sistema ortogonal* de vetores. Um vetor α é dito *normalizado* se $||\alpha|| = 1$. Um sistema ortogonal de vetores é dito um *sistema ortonormal* se cada vetor do sistema é normalizado.

Um sistema ortonormal de vetores $\{\xi_1,\ldots,\xi_n\}$ satisfaz

$$(\xi_k,\xi_l) = \delta_{kl} \quad , \quad k,l = 1,\ldots,n \ . \tag{C.25}$$

Qualquer conjunto ortonormal de vetores é linearmente independente porque se $\sum_k c_k\xi_k = 0$ então $0 = (\xi_l, \sum_k c_k\xi_k) = \sum_k c_k(\xi_l,\xi_k) = \sum_k c_k\delta_{kl} = c_l$. Assim, qualquer conjunto ortonormal de n vetores num espaço vetorial $\mathcal{V}$ de dimensão n é base de $\mathcal{V}$. Numa base ortonormal $\{\xi_1,\ldots,\xi_n\}$ um vetor α escreve-se $\alpha = \sum_j a_j\xi_j$, donde

$$(\xi_k,\alpha) = \sum_j (\xi_k, a_j\xi_j) = \sum_j a_j(\xi_k,\xi_j) = \sum_j a_j\delta_{kj} = a_k \ . \tag{C.26}$$

Portanto,

$$\alpha = \sum_k (\xi_k,\alpha)\,\xi_k \ . \tag{C.27}$$

Um cálculo análogo mostra que, em termos das componentes relativas a uma base ortonormal, o produto interno assume uma forma particularmente simples:

$$(\alpha, \beta) = \sum_k (\xi_k, \alpha)^* \, (\xi_k, \beta) \ . \tag{C.28}$$

Se $[A] = (a_{ij})$ é a matriz associada a um operador A numa base ortonormal $\{\xi_1, \ldots, \xi_n\}$, a Eq.(C.27), combinada com a definição C.3, fornece imediatamente

$$a_{ij} = (\xi_i, A\xi_j) \ . \tag{C.29}$$

Teorema C.4. Todo espaço vetorial de dimensão finita com produto interno possui uma base ortonormal.

Demonstração. Seja $\mathcal{V}$ um espaço vetorial com produto interno e seja $\{\alpha_1, \ldots, \alpha_n\}$ uma base de $\mathcal{V}$. A partir desta base pode-se obter uma base ortogonal pelo *processo de ortogonalização de Gram-Schmidt* (Hoffmann & Kunze 1971). Dividindo cada um dos vetores da referida base ortogonal por sua respectiva norma, obtém-se a base ortonormal procurada. □

C.5 Complemento Ortogonal e Soma Direta

Definição C.7. Seja $\mathcal{V}$ um espaço vetorial com produto interno e $\mathcal{S}$ um conjunto arbitrário de vetores de $\mathcal{V}$. O *complemento ortogonal* de $\mathcal{S}$ consiste no conjunto $\mathcal{S}^\perp$ de todos os vetores de $\mathcal{V}$ que são ortogonais a todo vetor de $\mathcal{S}$.

Definição C.8. O espaço vetorial $\mathcal{V}$ é a *soma direta* dos espaços vetoriais $\mathcal{U}$ e $\mathcal{W}$, e escrevemos $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$, se cada elemento α de $\mathcal{V}$ pode ser expresso de maneira *única* como $\alpha = \beta + \gamma$, onde $\beta \in \mathcal{U}$ e $\gamma \in \mathcal{W}$.

Teorema C.5. Se $\mathcal{W}$ é um subespaço de dimensão finita de um espaço vetorial $\mathcal{V}$ com produto interno, então $\mathcal{V} = \mathcal{W} \oplus \mathcal{W}^\perp$.

Demonstração. Seja $\{\xi_1, \ldots, \xi_n\}$ uma base ortonormal de $\mathcal{W}$, cuja existência é garantida pelo Teorema C.4. Qualquer vetor α de $\mathcal{V}$ pode ser escrito como

$$\alpha = \sum_{k=1}^{n} (\xi_k, \alpha) \, \xi_k \ + \ \gamma \tag{C.30}$$

onde

$$\gamma = \alpha - \sum_{k=1}^{n} (\xi_k, \alpha) \, \xi_k \ . \tag{C.31}$$

Com o uso de (C.25) verifica-se facilmente que $(\xi_i, \gamma) = 0$ para todo ξ_i. Segue-se que

γ é ortogonal a qualquer combinação linear de $\{\xi_1, \ldots, \xi_n\}$, portanto é um elemento de $\mathcal{W}^\perp$. De acordo com (C.30), temos $\alpha = \beta + \gamma$ com β em $\mathcal{W}$ e γ em $\mathcal{W}^\perp$. Resta provar que esta decomposição é única. Se $\alpha = \beta' + \gamma'$ com β' em $\mathcal{W}$ e γ' em $\mathcal{W}^\perp$, temos $\beta - \beta' = \gamma' - \gamma$. Como $\mathcal{W}^\perp$ também é um subespaço vetorial de $\mathcal{V}$ (verifique!), o vetor $\beta'' = \beta - \beta'$ pertence ao mesmo tempo a $\mathcal{W}$ e a $\mathcal{W}^\perp$, isto é, $(\beta'', \beta'') = 0$, o que implica $\beta'' = 0$. Logo $\beta' = \beta$ e $\gamma' = \gamma$, o que estabelece a unicidade e completa a demonstração do teorema. □

C.6 Adjunto de um Operador Linear

Uma noção importantíssima é a de adjunto $A^\dagger$ de um operador linear A. A fim de definir o adjunto e estabelecer suas propriedades essenciais, é necessário primeiro introduzir o conceito de *funcional linear*.

Definição C.9. Um *funcional linear* Φ num espaço vetorial $\mathcal{V}$ é uma regra que associa um escalar $\Phi(\alpha)$ a cada vetor α de $\mathcal{V}$ de tal modo que

$$\Phi(c_1\alpha_1 + c_2\alpha_2) = c_1\Phi(\alpha_1) + c_2\Phi(\alpha_2) \tag{C.32}$$

para todos os vetores α_1, α_2 e escalares c_1, c_2.

Por exemplo, se $\mathcal{V}$ é um espaço vetorial com produto interno e β é um vetor fixo, a função Φ_β definida por

$$\Phi_\beta(\alpha) = (\beta, \alpha) \tag{C.33}$$

é um funcional linear em virtude das propriedades do produto interno. Um resultado notável é que, se $\mathcal{V}$ tem dimensão finita, qualquer funcional linear é desta forma.

Teorema C.6. Seja $\mathcal{V}$ um espaço vetorial de dimensão finita com produto interno e Φ um funcional linear em $\mathcal{V}$. Então existe um único vetor β de $\mathcal{V}$ tal que $\Phi(\alpha) = (\beta, \alpha)$ para todo α de $\mathcal{V}$.

Demonstração. Seja $\{\alpha_1, \ldots, \alpha_n\}$ uma base ortonormal de $\mathcal{V}$ e defina

$$\beta = \sum_{j=1}^{n} \Phi(\alpha_j)^* \, \alpha_j \ . \tag{C.34}$$

Então

$$(\beta, \alpha_k) = \sum_{j=1}^{n} (\Phi(\alpha_j)^* \, \alpha_j, \alpha_k) = \sum_{j=1}^{n} \Phi(\alpha_j) \, (\alpha_j, \alpha_k) = \sum_{j=1}^{n} \Phi(\alpha_j) \, \delta_{jk} = \Phi(\alpha_k) \ , \tag{C.35}$$

e, para um vetor genérico α com $\alpha = \sum_k c_k \alpha_k$,

$$\Phi(\alpha) = \sum_{k=1}^{n} c_k \Phi(\alpha_k) = \sum_{k=1}^{n} c_k (\beta, \alpha_k) = (\beta, \sum_{k=1}^{n} c_k \alpha_k) = (\beta, \alpha) \ . \tag{C.36}$$

Para provar a unicidade, suponha que γ seja outro vetor tal que $\Phi(\alpha) = (\beta, \alpha) = (\gamma, \alpha)$ para todo α de $\mathcal{V}$. Então o vetor $\rho = \gamma - \beta$ é ortogonal a todos os vetores de $\mathcal{V}$, logo é ortogonal a si próprio. Assim, $(\rho, \rho) = 0$ donde $\rho = 0$ e $\gamma = \beta$. □

Teorema C.7. Se A é um operador linear num espaço vetorial $\mathcal{V}$ de dimensão finita com produto interno, existe um único operador linear $A^\dagger$ tal que[1]

$$(\beta, A\alpha) = (A^\dagger\beta, \alpha) \tag{C.37}$$

para todos os α, β em $\mathcal{V}$. O operador $A^\dagger$ é chamado de *adjunto* de A.

Demonstração. Como $\Phi(\alpha) = (\beta, A\alpha)$ é um funcional linear para β fixo, o Teorema C.6 nos diz que existe um único vetor β' em $\mathcal{V}$ tal que $(\beta, A\alpha) = (\beta', \alpha)$. O operador $A^\dagger$ é definido como a regra que associa β' a β:

$$A^\dagger\beta = \beta' \ . \tag{C.38}$$

Resta verificar que $A^\dagger$ é um operador linear. Sejam β, γ elementos de $\mathcal{V}$ e c_1, c_2 escalares. Para qualquer α,

$$\begin{aligned}(A^\dagger(c_1\beta + c_2\gamma), \alpha) &= (c_1\beta + c_2\gamma, A\alpha) = c_1^*(\beta, A\alpha) + c_2^*(\gamma, A\alpha) = c_1^*(A^\dagger\beta, \alpha) + c_2^*(A^\dagger\gamma, \alpha) \\ &= (c_1A^\dagger\beta, \alpha) + (c_2A^\dagger\gamma, \alpha) = (c_1A^\dagger\beta + c_2A^\dagger\gamma, \alpha) \ ,\end{aligned} \tag{C.39}$$

de modo que $A^\dagger(c_1\beta + c_2\gamma) = c_1A^\dagger\beta + c_2A^\dagger\gamma$ e $A^\dagger$ é linear. □

As seguintes propriedades valem num espaço vetorial de dimensão finita e são de fácil comprovação:

(1) $(A + B)^\dagger = A^\dagger + B^\dagger$;
(2) $(cA)^\dagger = c^*A^\dagger$ para $c \in \mathcal{F}$;
(3) $(AB)^\dagger = B^\dagger A^\dagger$;
(4) $(A^\dagger)^\dagger = A$.

C.7 Operadores Unitários e Auto-Adjuntos

Definição C.10. Um operador linear U num espaço vetorial $\mathcal{V}$ é dito *unitário* se o seu adjunto $U^\dagger$ coincide com o seu inverso U^{-1}, isto é,

$$UU^\dagger = U^\dagger U = I \ . \tag{C.40}$$

Teorema C.8. Se U é um operador linear num espaço vetorial de dimensão finita, qualquer uma das condições a seguir é necessária e suficiente para que U seja unitário: (i) $UU^\dagger = I$ ou $U^\dagger U = I$; (ii) $(U\xi, U\eta) = (\xi, \eta)$ para quaisquer vetores ξ, η de $\mathcal{V}$.

[1] Os matemáticos preferem denotar o complexo conjugado de um número complexo z por $\bar{z}$ e o adjunto de um operador linear A por A^*.

Demonstração. No que concerne à parte (i), se U é unitário tanto $UU^\dagger = I$ quanto $U^\dagger U = I$, e não há nada a demonstrar. Suponha, agora, que $U^\dagger U = I$. Pelo Teorema C.1 o operador U é necessariamente inversível uma vez que $U\alpha = 0$ implica $\alpha = I\alpha = U^\dagger U\alpha = U^\dagger 0 = 0$. Multiplicando $U^\dagger U = I$ pela esquerda por U e pela direita por U^{-1}, resulta $UU^\dagger = I$, e U é unitário. O mesmo argumento se aplica no caso $UU^\dagger = I$ intercambiando-se os papéis de U e $U^\dagger$, e a demonstração referente à parte (i) está completa. Quanto à parte (ii), se U é unitário então $(\xi,\eta) = (U^\dagger U\xi,\eta) = (U\xi,U\eta)$ pela definição de adjunto. Inversamente, se $(U\xi,U\eta) = (\xi,\eta)$ para quaisquer vetores ξ,η de $\mathcal{V}$, segue-se da definição de adjunto que $(U^\dagger U\xi,\eta) = (\xi,\eta)$ ou, equivalentemente, $((U^\dagger U - I)\xi,\eta) = 0$. Como ξ e η são vetores arbitrários, isto significa que $U^\dagger U = I$ e, pela parte (i), U é unitário. □

Os operadores unitários podem ser caracterizados pela propriedade de mapear bases ortonormais em bases ortonormais.

Teorema C.9. Se U é um operador unitário num espaço vetorial de dimensão finita $\mathcal{V}$ e $\{\alpha_1,\ldots,\alpha_n\}$ é base ortonormal de $\mathcal{V}$, então $\{U\alpha_1,\ldots,U\alpha_n\}$ também é base ortonormal de $\mathcal{V}$. Reciprocamente, se $\{\alpha_1,\ldots,\alpha_n\}$ e $\{\beta_1,\ldots,\beta_n\}$ são bases ortonormais de $\mathcal{V}$, existe um operador unitário U tal que $\beta_k = U\alpha_k$, $k = 1,\ldots,n$.

Demonstração. Se U é unitário, $(U\alpha_k,U\alpha_l) = (U^\dagger U\alpha_k,\alpha_l) = (\alpha_k,\alpha_l) = \delta_{kl}$ e $\{U\alpha_1,\ldots,U\alpha_n\}$ é base ortonormal de $\mathcal{V}$. Se $\{\alpha_1,\ldots,\alpha_n\}$ e $\{\beta_1,\ldots,\beta_n\}$ são bases ortonormais de $\mathcal{V}$, defina o operador linear U por $U\alpha_k = \beta_k$, $k = 1,\ldots,n$. Para quaisquer vetores $\xi = \sum_k x_k\,\alpha_k$ e $\eta = \sum_k y_k\,\alpha_k$ pertencentes a $\mathcal{V}$ temos

$$(U\xi,U\eta) = (\sum_{k=1}^{n} x_k\,U\alpha_k,\sum_{l=1}^{n} y_l\,U\alpha_l) = \sum_{k,l=1}^{n} x_k^*\,y_l\,(\beta_k,\beta_l) = \sum_{k,l=1}^{n} x_k^*\,y_l\,\delta_{kl}$$

$$= \sum_{k,l=1}^{n} x_k^*\,y_l\,(\alpha_k,\alpha_l) = (\sum_{k=1}^{n} x_k\,\alpha_k,\sum_{l=1}^{n} y_l\,\alpha_l) = (\xi,\eta) \qquad (C.41)$$

e U é unitário pelo Teorema C.8. □

Definição C.11. Um operador linear A num espaço vetorial $\mathcal{V}$ é dito *auto-adjunto* ou *hermitiano* se $A^\dagger = A$.

Os operadores unitários e auto-adjuntos são de importância fundamental para a Física, e seus autovalores e autovetores gozam de propriedades simples.

Teorema C.10. Se c é um autovalor de um operador unitário então $|c| = 1$. Os autovalores de um operador auto-adjunto são reais e os autovetores de um operador auto-adjunto associados a autovalores distintos são mutuamente ortogonais.

Demonstração. Se U é unitário e $U\alpha = c\,\alpha$ então $(\alpha,\alpha) = (U\alpha,U\alpha) = (c\,\alpha,c\,\alpha) = |c|^2(\alpha,\alpha)$. Logo, $|c|^2 = 1$ e, conseqüentemente, $|c| = 1$. Se A é auto-adjunto e $A\alpha = c\,\alpha$ então

$$c\,(\alpha,\alpha) = (\alpha,c\,\alpha) = (\alpha,A\alpha) = (A^\dagger\alpha,\alpha) = (A\alpha,\alpha) = (c\,\alpha,\alpha) = c^*\,(\alpha,\alpha)\ . \qquad (C.42)$$

Como $(\alpha,\alpha) \neq 0$, segue-se que $c = c^*$ e c é real. Finalmente, sejam α_1, α_2 vetores não-nulos tais que $A\alpha_1 = c_1\alpha_1$ e $A\alpha_2 = c_2\alpha_2$. Então

$$c_2(\alpha_1,\alpha_2) = (\alpha_1, c_2\alpha_2) = (\alpha_1, A\alpha_2) = (A^\dagger\alpha_1, \alpha_2) = (A\alpha_1,\alpha_2) = (c_1\alpha_1,\alpha_2) = c_1(\alpha_1,\alpha_2)\ , \tag{C.43}$$

onde usamos o fato de c_1 ser real. Da Eq.(C.43) deduz-se $(c_2 - c_1)(\alpha_1,\alpha_2) = 0$, de modo que $(\alpha_1,\alpha_2) = 0$ se $c_1 \neq c_2$. □

Teorema C.11. Todo operador auto-adjunto sobre um espaço vetorial de dimensão finita possui um autovetor.

Demonstração. Se o espaço vetorial é complexo, o resultado está contido no Teorema C.2. Se o espaço vetorial é real, o teorema fundamental da álgebra estabelece que existe um número complexo c tal que $\det([A] - c\boldsymbol{I}) = 0$ onde $[A]$ é a matriz do operador A em alguma base. Logo, existe uma matriz-coluna não-nula $[\alpha]$ tal que $[A][\alpha] = c[\alpha]$. Mas como A é auto-adjunto, c é real e podemos tomar $[\alpha]$ com elementos reais. Portanto, existe um vetor real não-nulo α tal que $A\alpha = c\alpha$. □

Um resultado crucial para aplicações à Física assegura que, dado um operador auto-adjunto A num espaço vetorial de dimensão finita $\mathcal{V}$, existe uma base de $\mathcal{V}$ constituída por autovetores de A. Como veremos, isto equivale a dizer que qualquer matriz hermitiana pode ser diagonalizada por uma transformação de similaridade efetuada por uma matriz unitária. Para provar este resultado em sua completa generalidade, precisamos de uma definição e de um teorema preliminar.

Definição C.12. Seja $\mathcal{V}$ um espaço vetorial e A um operador linear em $\mathcal{V}$. Se $\mathcal{W}$ é um subespaço vetorial de $\mathcal{V}$, dizemos que $\mathcal{W}$ é *invariante sob* A se para cada vetor α de $\mathcal{W}$ o vetor $A\alpha$ pertence a $\mathcal{W}$.

Teorema C.12. Seja $\mathcal{V}$ um espaço vetorial de dimensão finita com produto interno e seja A um operador linear em $\mathcal{V}$. Se $\mathcal{W}$ é um subespaço vetorial de $\mathcal{V}$ que é invariante sob A, então o complemento ortogonal de $\mathcal{W}$ é invariante sob $A^\dagger$.

Demonstração. Tomemos β pertencente a $\mathcal{W}^\perp$. Precisamos provar que $A^\dagger\beta$ está em $\mathcal{W}^\perp$, ou seja, que $(\alpha, A^\dagger\beta) = 0$ para todo α em $\mathcal{W}$. Mas se α está em $\mathcal{W}$ então $A\alpha$ também está em $\mathcal{W}$ e, portanto, $(A\alpha,\beta) = 0$, donde, pela definição de adjunto, $(\alpha, A^\dagger\beta) = (A\alpha,\beta) = 0$. □

Agora estamos prontos para enunciar e provar o resultado que constitui o ponto culminante deste apêndice: o teorema espectral para operadores auto-adjuntos.

Teorema C.13. Seja $\mathcal{V}$ um espaço vetorial de dimensão finita munido de produto interno e seja A um operador linear auto-adjunto em $\mathcal{V}$. Então existe uma base ortonormal de $\mathcal{V}$ constituída por autovetores de A.

Demonstração. Seja $N > 0$ a dimensão de $\mathcal{V}$. Pelo Teorema C.11, A possui um autovetor α, de modo que $\alpha_1 = \alpha/\parallel\alpha\parallel$ é um autovetor normalizado de A. Se $N = 1$, a demonstração está terminada. Procedemos, agora, por indução sobre N. Se $N > 1$, suponhamos que o teorema seja válido para espaços de dimensão menor que N.

Se $\mathcal{W}$ é o subespaço unidimensional gerado por α_1, dizer que α_1 é autovetor de A significa simplesmente que $\mathcal{W}$ é invariante sob A. Pelo Teorema C.12, o complemento ortogonal $\mathcal{W}^\perp$ é invariante sob $A^\dagger = A$. Mas $\mathcal{W}^\perp$, com o produto interno de $\mathcal{V}$, é um espaço vetorial com produto interno de dimensão $N-1$. Seja T o operador linear induzido sobre $\mathcal{W}^\perp$ por A, isto é, a restrição de A a $\mathcal{W}^\perp$. O operador T é bem definido sobre $\mathcal{W}^\perp$ porque $\mathcal{W}^\perp$ é invariante sob a ação de A. Claramente, T é auto-adjunto e, pela hipótese de indução, $\mathcal{W}^\perp$ possui uma base ortonormal $\{\alpha_2, \ldots, \alpha_n\}$ composta por autovetores de T, logo, de A. Como $\mathcal{V} = \mathcal{W} \oplus \mathcal{W}^\perp$, concluímos que $\{\alpha_1, \ldots, \alpha_n\}$ é a base procurada. □

Se A é um operador auto-adjunto, de (C.29) deduz-se que sua matriz $[A]$ numa base ortonormal $\{\alpha_1, \ldots, \alpha_n\}$ tem elementos

$$a_{ij} = (\alpha_i, A\alpha_j) = (A^\dagger \alpha_i, \alpha_j) = (A\alpha_i, \alpha_j) = (\alpha_j, A\alpha_i)^* = a_{ji}^* \ , \qquad (C.44)$$

de modo que

$$[A]^\dagger = [A] \qquad (C.45)$$

e $[A]$ é uma matriz hermitiana. Deste resultado inferimos que um operador linear é auto-adjunto se e somente se sua matriz em alguma (logo, toda) base ortonormal é hermitiana. Por outro lado, se U é um operador unitário,

$$\delta_{ij} = (\alpha_i, \alpha_j) = (U\alpha_i, U\alpha_j) = \sum_{k,l} u_{ki}^* u_{lj} (\alpha_k, \alpha_l) = \sum_{k,l} u_{ki}^* u_{lj} \delta_{kl} = \sum_k u_{ki}^* u_{kj} \ , \qquad (C.46)$$

donde

$$[U]^\dagger [U] = \boldsymbol{I} \ . \qquad (C.47)$$

Assim, $\det [U] \neq 0$ e $[U]$ é matriz inversível, de modo que

$$[U]^\dagger = [U]^{-1} \qquad (C.48)$$

e $[U]$ é uma matriz unitária. Segue-se que um operador linear é unitário se e somente se sua matriz em alguma (logo, toda) base ortonormal é unitária. É importante notar que num espaço vetorial sobre o corpo dos números reais uma matriz hermitiana é o mesmo que uma matriz real *simétrica*, ou seja, uma matriz que coincide com sua transposta. Da mesma forma, num espaço vetorial real uma matriz unitária reduz-se a uma matriz real *ortogonal*, isto é, uma matriz cuja transposta é igual à sua inversa.

De acordo com as observações acima, uma mudança de base ortonormal corresponde, por (C.16), a uma transformação de similaridade das matrizes associadas aos operadores realizada por uma matriz unitária, isto é,

$$[A]_{\mathcal{Y}} = [S]_{\mathcal{X}}^\dagger \, [A]_{\mathcal{X}} \, [S]_{\mathcal{X}} \ . \qquad (C.49)$$

Se $[A]$ é matriz hermitiana na base $\mathcal{X}$,

$$[A]_{\mathcal{Y}}^{\dagger} = \left([S]_{\mathcal{X}}^{\dagger}\,[A]_{\mathcal{X}}\,[S]_{\mathcal{X}}\right)^{\dagger} = [S]_{\mathcal{X}}^{\dagger}\,[A]_{\mathcal{X}}^{\dagger}\,[S]_{\mathcal{X}}^{\dagger\dagger} = [S]_{\mathcal{X}}^{\dagger}\,[A]_{\mathcal{X}}\,[S]_{\mathcal{X}} = [A]_{\mathcal{Y}} \;, \qquad (C.50)$$

de modo que $[A]$ também é hermitiana na base $\mathcal{Y}$. Analogamente, se a matriz $[U]$ é unitária na base $\mathcal{X}$,

$$[U]_{\mathcal{Y}}^{\dagger}[U]_{\mathcal{Y}} = [S]_{\mathcal{X}}^{\dagger}\,[U]_{\mathcal{X}}^{\dagger}\,[S]_{\mathcal{X}}\,[S]_{\mathcal{X}}^{\dagger}\,[U]_{\mathcal{X}}\,[S]_{\mathcal{X}} = [S]_{\mathcal{X}}^{\dagger}\,[U]_{\mathcal{X}}^{\dagger}\,[U]_{\mathcal{X}}\,[S]_{\mathcal{X}} = [S]_{\mathcal{X}}^{\dagger}\,[S]_{\mathcal{X}} = \boldsymbol{I} \qquad (C.51)$$

e $[U]$ também é matriz unitária na base $\mathcal{Y}$. Portanto, uma mudança de base ortonormal mapeia matrizes auto-adjuntas em matrizes auto-adjuntas e matrizes unitárias em matrizes unitárias. Notando que numa base ortonormal composta por autovetores de um operador linear A sua matriz associada $[A]$ é diagonal, as observações acima permitem uma tradução do Teorema (C.13) em termos de matrizes.

Teorema C.14. Qualquer matriz hermitiana pode ser diagonalizada por uma transformação de similaridade executada por uma matriz unitária. Em particular, uma matriz real simétrica pode ser diagonalizada por uma transformação de similaridade executada por uma matriz real ortogonal.

Apêndice D

Diferenciais Exatas

Neste apêndice definimos diferenciais exatas e estabelecemos as condições necessárias e suficientes para que uma dada forma diferencial seja exata.

Sejam $f_1, f_2, \ldots, f_n$ funções reais de n variáveis reais $x_1, x_2, \ldots, x_n$ e suponhamos que cada uma das funções f_i bem como cada uma das suas derivadas parciais $\partial f_i/\partial x_j$ sejam contínuas numa região aberta $\mathcal{O} \subset \boldsymbol{R}^n$. Restrições adicionais sobre a região $\mathcal{O}$ aparecerão de forma natural na discussão a ser desenvolvida abaixo. Sempre que revelar-se conveniente, utilizaremos a notação abreviada $x = (x_1, \ldots, x_n)$.

Definição D.1. A expressão

$$\Omega = \sum_{i=1}^{N} f_i \, dx_i \tag{D.1}$$

é dita uma diferencial exata em $\mathcal{O}$ se existe uma função $F(x_1, \ldots, x_n)$ tal que

$$dF = \sum_{i=1}^{N} f_i \, dx_i \tag{D.2}$$

em $\mathcal{O}$ ou, equivalentemente, se existe uma função $F(x_1, \ldots, x_n)$ tal que

$$f_i = \frac{\partial F}{\partial x_i} \tag{D.3}$$

em cada ponto da região $\mathcal{O}$.

Teorema D.1. A expressão $\Omega = \sum_i f_i \, dx_i$ é uma diferencial exata em $\mathcal{O}$ se e somente se

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i} \quad , \quad i, j = 1, \ldots, n \tag{D.4}$$

em todos os pontos de $\mathcal{O}$.

Demonstração. Provemos a necessidade de $(D.4)$. Se $\Omega = \sum_i f_i \, dx_i$ é uma diferencial exata em $\mathcal{O}$ existe uma função F tal que $f_i = \partial F/\partial x_i$. Da continuidade das derivadas parciais das funções f_i segue-se que

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial^2 F}{\partial x_j \partial x_i} = \frac{\partial^2 F}{\partial x_i \partial x_j} = \frac{\partial f_j}{\partial x_i} \quad , \quad i,j = 1,\ldots,n \tag{D.5}$$

em cada ponto de $\mathcal{O}$. Para demonstrar a suficiência, suponhamos a validade de $(D.4)$ e consideremos a função $F(x_1,\ldots,x_n)$ definida como a integral de linha de $\sum_i f_i dx_i$ ao longo de um segmento de reta unindo um ponto fixo P_0 ao ponto genérico $P = (x_1,\ldots,x_n)$ de $\mathcal{O}$. Escolhendo a origem do sistema de coordenadas no ponto P_0, as equações da linha reta em questão podem ser postas na forma paramétrica

$$\xi_i(\lambda) = \lambda\, x_i \quad , \quad \lambda \in [0,1] \ , \tag{D.6}$$

de modo que

$$F(x) = \int_{P_0}^{P} \sum_{k=1}^{N} f_k(\xi)\, d\xi_k = \sum_{k=1}^{N} \int_0^1 f_k(\lambda x)\, x_k\, d\lambda \ . \tag{D.7}$$

Conseqüentemente,

$$\begin{aligned} \frac{\partial F}{\partial x_i} &= \sum_{k=1}^{N} \int_0^1 \delta_{ik}\, f_k(\lambda x)\, d\lambda + \sum_{k=1}^{N} \int_0^1 \lambda \frac{\partial f_k}{\partial x_i}(\lambda x)\, x_k\, d\lambda \\ &= \int_0^1 f_i(\lambda x)\, d\lambda + \sum_{k=1}^{N} \int_0^1 \lambda\, x_k \frac{\partial f_i}{\partial x_k}(\lambda x)\, d\lambda \ , \end{aligned} \tag{D.8}$$

onde usamos $(D.4)$. Notando que, para x fixo,

$$\frac{df_i(\lambda x)}{d\lambda} = \sum_{k=1}^{N} \frac{\partial f_i}{\partial x_k}(\lambda x)\, x_k \ , \tag{D.9}$$

podemos escrever

$$\frac{\partial F}{\partial x_i} = \int_0^1 \Big[f_i(\lambda x) + \lambda \frac{df_i(\lambda x)}{d\lambda} \Big]\, d\lambda = \int_0^1 \frac{d}{d\lambda}\Big(\lambda f_i(\lambda x)\Big)\, d\lambda = \lambda f_i(\lambda x)\Big|_{\lambda=0}^{\lambda=1} = f_i(x) \ , \tag{D.10}$$

completando a demonstração do teorema. □

Note que o teorema se aplica a qualquer região aberta $\mathcal{O}$ dotada da seguinte propriedade: existe um ponto $P_0 \in \mathcal{O}$ tal que para qualquer ponto $P \in \mathcal{O}$ o segmento de reta de P_0 até P está inteiramente contido em $\mathcal{O}$. Um subconjunto aberto de $\boldsymbol{R}^n$ deste tipo é dito *estrelado* ("star-shaped") no ponto P_0 (Spivak 1965).

BIBLIOGRAFIA

Anderson, J. L. 1990 *Newton's First Two Laws of Motion Are Not Definitions*, American Journal of Physics **58**, 1192.

Aravind, P. K. 1989 *Geometrical Interpretation of the Simultaneous Diagonalization of Two Quadratic Forms*, American Journal of Physics **57**, 309.

Arfken, G. B. e Weber, **H. J.** 1995 *Mathematical Methods for Physicists* (Academic Press, New York).

Arnold, V. I. 1976 *Méthodes Mathématiques de la Mécanique Classique* (Éditions Mir, Moscou).

Arnowitt, R., Deser, S. e Misner, C. W. 1962 "The Dynamics of General Relativity" em Witten, L. (ed.) *Gravitation: An Introduction to Current Research* (Wiley, New York).

Barut, A. O. 1980 *Electrodynamics and Classical Theory of Fields and Particles* (Dover, New York).

Bateman, H. 1931 *On Dissipative Systems and Related Variational Principles*, Physical Review **38**, 815.

Bergmann, P. G. 1976 *Introduction to the Theory of Relativity* (Dover, New York).

Berry, M. V. 1978 "Regular and Irregular Motion" em Jorna, S. (ed.) *Topics in Nonlinear Dynamics*, American Institute of Physics Conference Proceedings **46**.

Cadoni, M., Carta, P. e Mignemi, S. 2000 *A Realization of the Infinite Dimensional Symmetries of Conformal Mechanics*, arXiv: hep-th/0004107.

Callen, H. 1960 *Thermodynamics* (Wiley, New York).

Cercignani, C. 1998 *Ludwig Boltzmann: The Man Who Trusted Atoms* (Oxford University Press, Oxford).

Chagas, E. F. das e Lemos, N. A. 1981 *Um Exemplo de Como Não Usar Teoremas Matemáticos em Problemas Físicos*, Revista Brasileira de Física **11**, 481.

Chandre, C. e Jauslin, H. R. 1998 *A Version of Thirring's Approach to the Kolmogorov-Arnold-Moser Theorem for Quadratic Hamiltonians with Degenerate Twist*, Journal of Mathematical Physics **39**, 5856.

Christ, N. H. e Lee, T. D. 1980 *Operator Ordering and Feynman Rules in Gauge Theories*, Physical Review **D22**, 939.

Corben, H. C. e Stehle, P. 1960 *Classical Mechanics* (Dover, New York).

Courant, R. e Hilbert, D. 1953 *Methods of Mathematical Physics*, Vol. I (Interscience Publishers, New York).

Crawford, F. S. 1990 *Elementary Examples of Adiabatic Invariance*, American Journal of Physics **58**, 337.

Deser, S., Jackiw, R. e Templeton, S. 1982 *Topologically Massive Gauge Theories*, Annals of Physics **140**, 372.

Desloge, E. A. 1982 *Classical Mechanics*, 2 Vols. (Robert E. Krieger, Malabar, FL).

Dettman, J. W. 1986 *Introduction to Linear Algebra and Differential Equations* (Dover, New York).

Dirac, P. A. M. 1964 *Lectures on Quantum Mechanics* (Yeshiva University, New York).

Dugas, R. 1988 *A History of Mechanics* (Dover, New York).

Eisenbud, L. 1958 *On the Classical Laws of Motion*, American Journal of Physics **26**, 144.

Elsgoltz, L. 1969 *Ecuaciones Diferenciales y Cálculo Variacional* (Editorial Mir, Moscou).

Epstein, S. T. 1982 *The Angular Velocity of a Rotating Rigid Body*, American Journal of Physics **50**, 948.

Farina de Souza, C. e Gandelman, M. M. 1990 *An Algebraic Approach for Solving Mechanical Problems*, American Journal of Physics **58**, 491.

Fetter, A. L. e Walecka, J. D. 1980 *Theoretical Mechanics of Particles and Continua* (McGraw-Hill, New York).

Feynman, R. P., Leighton, R. B. e Sands, M. 1963 *The Feynman Lectures on Physics*, Vols. I–III (Addison-Wesley, Reading, MA).

Feynman, R. P. e Hibbs, A. 1965 *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York).

Feynman, R. P. 1985 *Surely You're Joking, Mr. Feynman!* (Bantam Books, New York).

Gantmacher, F. 1970 *Lectures in Analytical Mechanics* (Mir Publishers, Moscow).

Gelfand, I. M. e Fomin, S. V. 1963 *Calculus of Variations* (Prentice-Hall, Englewood Cliffs, NJ).

Gleiser, R. J. e Kozameh, C. N. 1980 *A Simple Application of Adiabatic Invariance*, American Journal of Physics **48**, 756.

Goldstein, H. 1980 *Classical Mechanics* (Addison-Wesley, Reading, MA), 2ª edição.

Gradshteyn, I. S. e Ryzhik, I. M. 1980 *Tables of Integrals, Series and Products* (Academic Press, New York).

Halmos, P. R. 1974 *Finite-Dimensional Vector Spaces* (Springer, New York).

Hamermesh, M. 1962 *Group Theory and Its Application to Physical Problems* (Dover, New York).

Hand, L. N. e Finch, J. D. 1998 *Analytical Mechanics* (Cambridge University Press, Cambridge).

Hartmann, H. 1972 *Die Bewegung eines Körpers in einem ringförmigen Potentialfeld*, Theoretica Chimica Acta **24**, 201.

Havas, P. 1957 *The Range of Application of the Lagrange Formalism - I*, Nuovo Cimento Suppl. **5**, 363.

Henneaux, M. e Teitelboim, C. 1992 *Quantization of Gauge Systems* (Princeton University Press, New Jersey).

Hoffmann, K. e Kunze, R. 1961 *Álgebra Linear* (Editora Polígono, São Paulo).

Huang, K. 1963 *Statistical Mechanics* (Wiley, New York).

Jackson, J. D. 1999 *Classical Electrodynamics* (Wiley, New York), 3ª edição.

José, J. V. e Saletan, E. J. 1998 *Classical Dynamics: A Contemporary Approach* (Cambridge University Press, Cambridge).

Konopinski, E. J. 1969 *Classical Descriptions of Motion* (W. H. Freeman & Co., San Francisco).

Lanczos, C. 1970 *The Variational Principles of Mechanics* (Dover, New York).

Landau, L. D. e Lifchitz, E. 1966 *Mécanique* (Éditions Mir, Moscou).

Lee, T. D. 1981 *Particle Physics and Introduction to Field Theory* (Harwood Academic Publishers, New York).

Lemos, N. A. 1979 *Canonical Approach to the Damped Harmonic Oscillator*, American Journal of Physics **47**, 857.

Lemos, N. A. e Natividade, C. P. 1987 *Harmonic Oscillator in Expanding Universes*, Nuovo Cimento B **99**, 211.

Lemos, N. A. 1991 *Remark on Rayleigh's Dissipation Function*, American Journal of Physics **59**, 660.

Lemos, N. A. 1993 *Symmetries, Noether's Theorem and Inequivalent Lagrangians Applied to Nonconservative Systems*, Revista Mexicana de Física **39**, 304.

Lemos, N. A. 1996 *Singularities in a Scalar Field Quantum Cosmology*, Physical Review **D53**, 4275.

Lemos, N. A. 2000a *Short Proof of Jacobi's Identity for Poisson Brackets*, American Journal of Physics **68**, 88.

Lemos, N. A. 2000b *Uniqueness of the Angular Velocity of a Rigid Body: Correction of Two Faulty Proofs*, American Journal of Physics **68**, 668.

Lemos, N. A. 2003 *Sutilezas dos Vínculos Não-Holônomos*, Revista Brasileira de Ensino de Física **25**, 28.

Lemos, N. A. 2005 *Formulação Geométrica do Princípio de d'Alembert*, Revista Brasileira de Ensino de Física **27**, 483.

Leubner, C. 1979 *Coordinate-Free Rotation Operator*, American Journal of Physics **47**, 727.

Lewis Jr., H. R. e Riesenfeld, W. B. 1969 *An Exact Quantum Theory of the Time-Dependent Harmonic Oscillator and of a Charged Particle in a Time-Dependent Electromagnetic Field*, Journal of Mathematical Physics **10**, 1458.

Lissauer, J. L. 1999 *Chaotic Motion in the Solar System*, Reviews of Modern Physics **71**, 835.

Lorentz, H. A., Einstein, A., Minkowski, H. e Weyl, H. 1952 *The Principle of Relativity* (Dover, New York).

Mann, R. A. 1974 *The Classical Dynamics of Particles* (Academic Press, New York).

Marion, J. B. e Heald, M. A. 1980 *Classical Electromagnetic Radiation* (Academic Press, New York), 2ª edição.

Marion, J. B. e Thornton, S. T. 1995 *Classical Dynamics of Particles and Systems* (Saunders College Publishing, Fort Worth), 4ª edição.

Maxwell, J. C. 1891 *A Treatise on Electricity and Magnetism*, 2 Vols., 3ª edição, reimpressa em 1954 (Dover, New York).

Mehra, J. 1994 *The Beat of a Different Drum* (Oxford University Press, Oxford).

Neĭmark, J. I. e Fufaev, N. A. 1972 *Dynamics of Nonholonomic Systems* (American Mathematical Society, Providence).

Noether, E. 1918 *Invariante Variationsprobleme*, Nachr. Gesell. Wissenchaft. Göttingen **2**, 235. [Tradução para o inglês: Transp. Theory Stat. Phys. **1**, 186 (1971).]

Núñes-Yepes, H. N., Delgado, J. e Salas-Brito, A. L. 2001 *Variational Equations of Lagrangian Systems and Hamilton's Principle*, arXiv: math-ph/0107006.

Ohanian, H. C. 1976 *Gravitation and Spacetime* (W. W. Norton & Company, New York).

Ott, E. 1993 *Chaos in Dynamical Systems* (Cambridge University Press, Cambridge).

Ozorio de Almeida, A. M. 1995 *Sistemas Hamiltonianos: Caos e Quantização* (Editora da Unicamp, Campinas).

Pais, A. 1982 *'Subtle is the Lord ...' The Science and the Life of Albert Einstein* (Oxford University Press, Oxford).

Pars, L. 1965 *A Treatise on Analytical Dynamics* (Ox Bow Press, Woodbridge, CT).

Pathria, R. K. 1972 *Statistical Mechanics* (Pergamon Press, Oxford).

Pauli, W. 1958 *Theory of Relativity* (Dover, New York).

Pearlman, N. 1967 *Vector Representation of Rigid Body Rotation*, American Journal of Physics **35**, 1164.

Protter, M. H. e Morrey Jr., C. B. 1985 *Intermediate Calculus* (Springer, New York).

Rindler, W. 1982 *Introduction to Special Relativity* (Clarendon Press, Oxford).

Romer, R. H. 1978 *Demonstration of the Intermediate-Axis Theorem*, American Journal of Physics **46**, 575.

Rosen, G. 1969 *Formulations of Classical and Quantum Dynamical Theory* (Academic Press, New York).

Rund, H. 1966 *The Hamilton-Jacobi Theory in the Calculus of Variations* (D. Van Nostrand, London).

Russell, J. S. 1844 *Report on Waves*, 14th Meeting of the British Association for the Advancement of Science, 311 (John Murray, London).

Saletan, E. J. e Cromer, A. H. 1970 *A Variational Principle for Nonholonomic Systems*, American Journal of Physics **38**, 892.

Saletan, E. J. e Cromer, A. H. 1971 *Theoretical Mechanics* (Wiley, New York).

Scheck, F. 1994 *Mechanics – From Newton's Laws to Deterministic Chaos* (Springer, Berlin).

Schiff, L. I. 1968 *Quantum Mechanics* (McGraw-Hill, New York), 3ª edição.

Schrödinger, E. 1982 *Collected Papers on Wave Mechanics* (Chelsea Publishing Co., New York). Tradução francesa: *Mémoires sur la Mécanique Ondulatoire* (Librairie Félix Alcan, Paris, 1933).

Sharan, P. 1982 *Two Theorems in Classical Mechanics*, American Journal of Physics **50**, 351.

Shinbrot, T., Grebogi, G., Wisdom, J. e Yorke, J. A. 1992 *Chaos in a Double Pendulum*, American Journal of Physics **60**, 491.

Sivardière, J. 1983 *A Simple Mechanical Model Exhibiting a Spontaneous Symmetry Breaking*, American Journal of Physics **51**, 1016.

Sivardière, J. 1986 *Using the Virial Theorem*, American Journal of Physics **54**, 1100.

Sommerfeld, A. 1952 *Mechanics* (Academic Press, New York).

Soper, D. E. 1976 *Classical Field Theory* (Wiley, New York).

Spiegel, M. R. 1971 *Cálculo Avançado* (McGraw-Hill do Brasil, Rio de Janeiro, RJ)

Spivak, M. 1965 *Calculus on Manifolds* (Benjamin, Menlo Park, CA).

Stadler, W. 1982 *Inadequacy of the Usual Newtonian Formulation for Certain Problems in Particle Mechanics*, American Journal of Physics **50**, 595.

Sternberg, S. 1994 *Group Theory and Physics* (Cambridge University Press, Cambridge).

Sundermeyer, K. 1982 *Constrained Dynamics* (Springer, New York).

Symon, K. R. 1971 *Mechanics* (Addison-Wesley, Reading, MA), 3ª edição.

Synge, J. L. e Griffith, B. A. 1959 *Principles of Mechanics* (McGraw-Hill, New York), 3ª edição.

Tabor, M. 1989 *Chaos and Integrability in Nonlinear Dynamics* (Wiley, New York).

Thirring, W. 1997 *Classical Mathematical Physics* (Springer, Berlin).

Tiersten, M. S. 1991 *Moments Not to Forget — The Conditions For Equating Torque and Rate of Change of Angular Momentum Around the Instantaneous Center*, American Journal of Physics **59**, 733.

Tufillaro, N. 1986 *Integrable Motion of a Swinging Atwood's Machine*, American Journal of Physics **54**, 142.

Tung, W. 1985 *Group Theory in Physics* (World Scientific, Singapore).

van Dam, N. e Wigner, E. 1966 *Instantaneous and Asymptotic Conservation Laws for Classical Relativistic Mechanics of Interacting Point Particles*, Physical Review **142**, 838.

van der Waerden, B. L. (ed.) 1967 *Sources of Quantum Mechanics* (Dover, New York).

Weinstock, R. 1961 *Laws of Classical Motion: What's* $\mathbf{F}$*? What's* m*? What's* $\mathbf{a}$*?*, American Journal of Physics **29**, 698.

Weinstock, R. 1974 *Calculus of Variations* (Dover, New York).

Whittaker, E. T. 1944 *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies* (Dover, New York).

Whittaker, E. T. 1951 *A History of the Theories of Aether and Electricity*, Vol. I: *The Classical Theories* (Thomas Nelson and Sons Ltd., London).

Wisdom, J. 1987 "Chaotic Behaviour in the Solar System" em Berry, M., Percival, I. C. e Weiss, N. O. (eds.) *Dynamical Chaos* (The Royal Society, London).

Yourgrau, W. e Mandelstam, S. 1968 *Variational Principles in Dynamics and Quantum Theory* (Dover, New York).

Índice Remissivo

Ação 51
  como fase de processo ondulatório 312
  como função das coordenadas 292-4
  reduzida 232
Ação à distância na relatividade especial 205-7
Aceleração da gravidade efetiva 205
Adjunto de um operador linear 366
Álgebra de Lie de funções no espaço de fase 253
Algoritmo de Dirac-Bergmann 268
Analogia entre mecânica clássica e óptica geométrica 311
Ângulos de Euler 91
Anomalia excêntrica (Problema 1.16) 42
Antikink 339
Autovalor e autovetor 362

Base ortonormal 365
Braquistócrona 49

Cálculo das variações 43
  equação de Euler 46
  lema fundamental 46
  problema mais simples 44
  problema isoperimétrico (Problema 2.5) 76
Campo escalar 189
  tensorial 189
  de Schrödinger como sistema hamiltoniano com vínculos 341
  eletromagnético como sistema hamiltoniano com vínculos 344
Carga de Noether 336
Catenária 49
Ciclóide 49-50
Colisões relativísticas 200
Complemento ortogonal 365
Condição de Noether 72
Condições de consistência 266
Cone de luz 183
Cone do corpo 136
  do espaço 136
Conjunto completo de observáveis comutativos (Problema 9.18) 318
Constante de movimento 63
Construção de Poinsot 134
Convenção de soma sobre índices repetidos de Einstein 179
Coordenada cíclica 64
Coordenadas generalizadas 19
  parabólicas (Problema 9.2) 315
  normais 163
Corpo rígido, deslocamentos possíveis 90
  número de graus de liberdade 83
Correspondência entre vetores e matrizes anti-simétricas 95
Cosmologia quântica 230
Cossenos diretores 84

Delta de Kronecker 353
Densidade hamiltoniana 330
Densidade lagrangiana 319
Derivada funcional 327
  regra da cadeia 327
Derivada parcial total 334
Desigualdade de Schwarz 363
Desigualdade triangular 364
Deslocamentos virtuais 14
Deslocamentos reais 14
Díada 114
Diádico 114
Diagonalização do tensor de inércia 122
Diagonalização simultânea de duas formas quadráticas 163-166
Diagrama de Minkowski 184
Difeomorfismo 25
Diferencial exata 373
Dilatação espacial 74

Divergência no espaço de Minkowski 190-191

Efeito Compton 202
Eixo instantâneo de rotação 96
Eixos principais de inércia 123
de cubo homogêneo em relação a um vértice 126-128
de placa triangular homogênea 124
Eletrodinâmica de Weber (Problemas 1.9 e 7.14) 41, 237
Elipsóide de inércia 135
Energia cinética de um corpo rígido 119
Energia, lei de conservação da 10, 68
e invariância sob translações temporais 70
Energia relativística 198
Energia de repouso 198
Equação de Euler 46
Equação de Hamilton-Jacobi 281
separação de variáveis 287
integral (solução) completa 283
Equação de Klein-Gordon 325
de *sine*-Gordon (Problema 10.8) 348
Equação de Schrödinger e teoria de Hamilton-Jacobi 313
Equações de Hamilton 215
covariantes 223
em teoria de campos 331
forma simplética 247
Equações de Lagrange 24
invariância 25
para campos 324, 329
Equações de Maxwell 192
em forma covariante 207
Equações de movimento de Euler para um corpo rígido 131
Escalar de Lorentz 179
Espaço de configuração 20
Espaço de fase 215
estendido 229
reduzido 228
Espaço de Minkowski 179
Espectro de um operador linear 362
Estabilidade, critério de 150
Estabilidade da rotação uniforme 134

Fator integrante 13
Fixação do calibre 269
Força 6
aplicada 17
central 33
centrífuga 102
conservativa 10
de atrito viscoso 31
de Coriolis 102
de Euler 102
de Lorentz 30, 197
de Minkowski 197
de vínculo 17
generalizada 20
generalizada de vínculo 56
Forma quadrática 157
Formalismo ADM 230
Fóton 202
Freqüências características 159
positividade 170
Freqüências comensuráveis 296
fundamentais 298
racionalmente independentes 304
Função
de dissipação de Rayleigh 41
de primeira classe 270
de segunda classe 270
energia 69
geradora de transformação canônica 241
homogênea 357
principal de Hamilton 292
Funcional 44
linear 366

Geodésicas da esfera 47
do cone (Problema 2.1) 76
no espaço de configuração 234
Gradiente no espaço de Minkowski 187
Gram-Schmidt, processo de ortogonalização de 365
Graus de liberdade 20
Grupo 89-89
das matrizes ortogonais 90
das transformações canônicas 250
de Lorentz 181
restrito 182
setores 182
de rotações 90
geradores infinitesimais 100
álgebra de Lie 100

constantes de estrutura 100
lei de composição 90
real simplético 250

Hamiltoniana 214
como geradora da evolução temporal 254
como momento conjugado ao tempo 226
conexão com a energia total 218
conservação 218
de partícula em campo eletromagnético externo 217
relativística 222
de partícula em campo eletromagnético externo 204
total 269
Herpolodia 136
Hipótese ergódica 303
adiabática 311

Identidade de Jacobi 253
demonstração 255
Igualdade forte 266
fraca 266
Índice livre 352
mudo 351
Integrabilidade 300, 306
Integral de Jacobi 69
Integrais de trajetória 294
Integrável, sistema 302
Intervalo invariante 179
do tipo espaço 183
do tipo tempo 183
do tipo luz 184
Invariância rotacional e conservação do momento angular 68
Invariância sob translações temporais e conservação da energia 70
Invariância translacional e conservação do momento linear 67
Invariante adiabático 308
Involução, constantes de movimento em 300

Kepler, equação de (Problema 1.16) 42
lei das áreas de 33
problema de 34
problema relativístico de (Problema 7.16) 237
terceira lei de 35
Kink 339

Lacunas de Kirkwood 306
Lagrangiana 24
covariante de partícula num campo eletromagnético externo 204
da corda vibrante 321
de Bateman (Problema 1.14) 42
de Chern-Simons (Problema 10.12) 349
de Klein-Gordon 325
de partícula num campo eletromagnético externo 31
de partícula num potencial central 33
de Proca (Problema 10.11) 349
de *sine*-Gordon (Problema 10.8) 348
do campo de Schrödinger 325
do campo eletromagnético 326
do pêndulo duplo 28, 29
em teoria de campos 319
relativística 203
de partícula num campo eletromagnético externo 203
Lagrangianas equivalentes 52
Lei da conservação local da carga elétrica 192
Lei de transformação das componentes de um tensor 117
Leis físicas em forma manifestamente covariante 191, 192
Libração 296
Limiar energético 193
Limite contínuo de sistema discreto 321
Linha de universo 184

Máquina de Atwood 18
oscilante (Problema 1.5) 39
e o método de Hamilton-Jacobi (Problema 9.12) 317
Matriz de inércia 119
Matriz
hermitiana 370
hessiana 214
modal 164
ortogonal 87
simplética 248
unitária 370

Mecânica conforme (Problema 8.12) 276
Mecânica generalizada (Problema 2.13) 78
Mecânica ondulatória 311
Média, princípio da 309
teorema da 310
Métrica 180
Modos normais de vibração 158
ortogonalidade 171
completeza 171
Moeda rolante 128-130
Molécula triatômica linear 167
Momento angular 8
do corpo rígido 111-113
lei de conservação 9
e invariância rotacional 68
Momento (canônico) conjugado 64
Momento linear 7
lei de conservação 8
e invariância translacional 67
Momento linear relativístico 196
Momentos de inércia 119
principais de inércia 123
Monopolo magnético (Problema 1.15) 42
Movimento caótico 304
regular 303
Movimento hiperbólico 199
Mudança de base 361
Multiperiódicos, sistemas 295
Multiplicadores de Lagrange 55
e forças generalizadas de vínculo 56
como coordenadas generalizadas 62

Newton, leis de 6
Norma 363
canônica 363
Nutação 141

Onda solitária 337
Operador
auto-adjunto ou hermitiano 368
de d'Alembert ou d'Alembertiano 191
de rotação finita 115
linear 359
inversível 359
unitário 367
Oscilador harmônico amortecido (Problemas 1.14 e 2.14) 42, 79
e a teoria de Hamilton-Jacobi 286

Parêntese de Dirac 267
Parênteses de Lagrange 245
e critério de canonicidade 245
invariância sob transformações canônicas 249
Parênteses de Poisson 250
do momento angular 260
invariância sob transformações canônicas 251
e critério de canonicidade 252
fundamentais 253
para campos 331
fundamentais 332
propriedades algébricas 253
Partícula de massa zero 200
Patinete 58
Pêndulo
amortecido 32
cicloidal (Problema 1.4) 39
de Foucault 105
duplo plano 11-12
equações de Lagrange 26, 29
lagrangiana 28
elástico (Problema 1.10) 41
esférico (Problema 1.2) 39
Pêndulos acoplados 160
Pequenas oscilações em torno de movimento estacionário 154
Pião simétrico 138
estabilidade 143
nutação e precessão 144
Polodia 135
Ponto de vista ativo 88
passivo 88
Ponto de equilíbrio 150
Potencial generalizado 29
efetivo 140, 155
de Hartmann 291
Precessão 144
Princípio
de d'Alembert 14-15
de Hamilton (da mínima ação) 52
no espaço de fase 224
de Jacobi 234
de Maupertuis 231
e princípio de Fermat 233
estendido a sistemas não-conservativos (Problema 7.18) 238
Princípio dos trabalhos virtuais 16

Problema de N corpos (Problema 7.6) 235
isoperimétrico (Problema 2.5) 75
Produto interno 363
canônico 363-364
Produto tensorial (diádico) 113
escalar no espaço de Minkowski 186-187
Produtos de inércia 118

Quantização canônica 258
Quadraturas, solução por 139
Quadriaceleração 194
Quadrimomento 194-195
Quadrivelocidade 194
Quadrivetor contravariante 186
Quebra espontânea de simetria 153
covariante 187
densidade de carga-corrente 192

Referenciais inerciais 6
Regras de quantização de Wilson-Sommerfeld ou Bohr-Sommerfeld 311
Rotação finita 115
Rotações infinitesimais 83, 94
comutatividade 94
Rotações no plano 87
Rotacional no espaço de Minkowski 191

Separatriz 296-297
Simetrias e eixos principais de inércia 126
e leis de conservação 63, 75, 217-219
Sistema com o tempo parametrizado 228
e partícula livre relativística 228
Sistema hamiltoniano
integrável 302
multiperiódico 295
separável 295
Sistema ortonormal de vetores 364
Sóliton 341
Solução completa 313
Solução das equações de movimento em série de parênteses de Poisson 257
Soma direta 365
Subespaço invariante 369
Superfície mínima de revolução 48
Super-hamiltoniana 226

Taxa de variação de um vetor 95-97
Tempo como variável canônica 226
Tempo próprio 194
Tensor de energia-momento 336
do campo eletromagnético (Problemas 6.18 e 10.6) 210, 347
do campo escalar 337
Tensor
de inércia 113
em notação diádica 118
do campo eletromagnético 193
dual (Problema 6.17) 209-210
identidade (unidade) 114
métrico 180
momento angular (Problemas 10.5 e 10.13) 346, 350
Teorema
da função implícita 214
da recorrência de Poincaré 263
e segunda lei da termodinâmica 264
de Bertrand 35
de Chasles 91
de Euler das funções homogêneas 357
de Euler sobre os movimentos possíveis de um corpo rígido 89-90
de Jacobi 283
de Kolmogorov-Arnold-Moser (KAM) 305
de Liouville 262
de Liouville-Arnold 302
de Noether 71
em teoria de campos 333
generalizado 73
de Poisson 256
de Stäckel 289
de Wigner-Van Dam 207
do eixo perpendicular 122
do virial 219
e equação de estado do gás ideal 221
dos eixos paralelos (Steiner) 121
espectral 369
Toro invariante 303-304
ressonante 305
não-ressonante 305
Trabalho virtual 18
Transformação canônica 239-252
e evolução temporal 254-255
Transformação canônica infinitesimal 253
Transformação
de calibre (Problema 2.11) 78
de primeira espécie ou global 337

e conservação da carga elétrica 337
local ou de segunda espécie (Problema 10.10) 348
de Galileu (Problema 2.15) 79
de Legendre 214
de Lorentz 177
imprópria 182
infinitesimal (Problema 6.5) 208
ortócrona 183
própria 182
pura 182
restrita 182
de ponto 25
de similaridade 362
infinitesimal 65
e teorema de Noether 80
linear 359
ortogonal 87
Translação rígida 66

Unicidade do vetor velocidade angular 97

Variação de um funcional 50
de uma função 50
Variáveis de ação-ângulo 297
no problema de Kepler 306
Variedade de configuração 20
Variedade diferenciável 20
Velha teoria quântica 310
Velocidade de grupo (Problema 9.17) 318
Velocidade angular 96
expressa em termos dos ângulos de Euler 98-99
definição analítica (Problema 3.7) 109
Vetor associado a uma rotação infinitesimal 95
Vetor deslocamento angular, inexistência de (Problema 3.6) 108
Vetor de Laplace-Runge-Lenz (Problema 2.22) 80
Vínculo da lei de Gauss 344
Vínculos 11
de primeira classe 270
de segunda classe 270
holônomos 12
incluídos na lagrangiana 62
ideais 16
não-holônomos 12
Vínculos (*continuação*)
não-integráveis 13
primários 268
secundários 268
Virial de Clausius 219

GRÁFICA PAYM
Tel. (011) 4392-3344
paym@terra.com.br

da do corpo de conhecimentos que constitui o cerne tradicional da mecânica analítica, tópicos de maior atualidade, embora mais difíceis, tais como os sistemas hamiltonianos com vínculos, os sistemas integráveis e o teorema de Kolmogorov-Arnold-Moser, recebem um tratamento introdutório que visa despertar a curiosidade e estimular estudos mais avançados.

Nivaldo Agostinho Lemos é professor de Física da Universidade Federal Fluminense, em Niterói, RJ. Obteve seu doutorado em 1981 no Centro Brasileiro de Pesquisas Físicas e realizou pós-doutorado no Massachusetts Institute of Technology (1994-1996). Sua principal área de pesquisa é a cosmologia quântica, com algumas incursões pela teoria quântica de campos. É autor de diversas publicações em periódicos científicos internacionais, bem como em revistas dedicadas ao ensino de Física.

Editora Livraria da Física
www.livrariadafisica.com.br

Indispensável para a compreensão das teorias físicas contemporâneas, a mecânica analítica cresceu em importância devido ao desenvolvimento recente das teorias do caos e dos sistemas dinâmicos. Neste livro, destinado primordialmente a cursos de graduação, os formalismos lagrangiano e hamiltoniano são introduzidos de forma clara, com ênfase nas propriedades de simetria e invariância e nos aspectos estruturais da mecânica.

Numerosos exemplos tornam o texto acessível aos estudantes típicos. O conteúdo tradicional da mecânica analítica é exposto de forma ampla e detalhada. Tópicos de maior atualidade, tais como os sistemas hamiltonianos com vínculos, os sistemas integráveis e o teorema de Kolmogorov-Arnold-Moser, recebem um tratamento introdutório que visa despertar a curiosidade e estimular estudos mais avançados.

Livraria da Física Editora

Editora Livraria da Física
www.livrariadafisica.com.br

ISBN 85-88325-24-1

9 788588 325241